

The Assessment of Habitat Condition and Conservation status of Lowland British Woodlands using Earth Observation Techniques

MATTHEW SUMNALL

A thesis submitted in partial fulfilment of the requirements of Bournemouth
University for the degree of Doctor of Philosophy

May 2013

Bournemouth University

In collaboration with

Centre for Ecology and Hydrology

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Abstract

The successful implementation of habitat preservation and management demands regular and spatially explicit monitoring of conservation status at a range of scales based on indicators. Woodland condition can be described in terms of compositional and structural attributes (e.g. overstorey, understorey, ground flora), evidence of natural turnover (e.g. deadwood and tree regeneration), and anthropogenic influences (e.g. disturbance, damage). Woodland condition assessments are currently conducted via fieldwork, which is hampered by cost, spatial coverage, objectiveness and repeatability. This project evaluates the ability of airborne remote sensing (RS) techniques to assess woodland condition, utilising a sensor-fusion approach to survey a forest study site and develop condition indicators. Here condition is based on measures of structural and compositional diversity in the woodland vertical profile, with consideration of the presence of native species, deadwood, and tree regeneration.

A 22 km² study area was established in the New Forest, Hampshire, UK, which contained a variety of forest types, including managed plantation, semi-ancient coniferous and deciduous woodland. Fieldwork was conducted in 41 field plots located across this range of forest types, each with varying properties. The field plots were 30x30m in size and recorded a total of 39 forest metrics relating to individual elements of condition as identified in the literature.

Airborne hyperspectral data (visible and near-infrared) and small footprint LiDAR capturing both discrete-return (DR) and full-waveform (FW) data were acquired simultaneously, under both leaf-on and leaf-off conditions in 2010. For the combined leaf-on and leaf-off datasets a total of 154 metrics were extracted from the hyperspectral data, 187 metrics from the DR LiDAR and 252 metrics from the FW LiDAR. This comprised both area-based and individual tree crown metrics. These metrics were entered into two statistical approaches, ordinary least squares and Akaike information criterion regression, in order to estimate each of the 39 field plot-level forest variables. These estimated variables were then used as inputs to six forest condition assessment approaches identified in the literature.

In total, 35 of the 39 field plot-level forest variables could be estimated with a validated NRMSE value below 0.4 using RS data (23 of these models had NRMSE values below 0.3). Over half of these models involved the use of FW LiDAR data on its own or combined with hyperspectral data, demonstrating this to be single most able dataset. Due to the synoptic coverage of the RS data, each of these field plot variables could be estimated and mapped continuously over the entire study site at the 30x30m resolution (i.e. field plot-level scale).

The RS estimated field variables were then used as inputs to six forest condition assessment approaches identified in the literature. Three of the derived condition indices were successful based on correspondence with field validation data and woodland compartment boundaries. The three successful condition assessment methods were driven primarily by tree size and tree size variation. The best technique for assessing woodland condition was a score-based method which combined seventeen inputs which relate to tree species composition, tree size and variability, deadwood, and understory components; all of which were shown to be derived successfully from the appropriate combination of airborne hyperspectral and LiDAR datasets.

The approach demonstrated in this project therefore shows that conventional methods of assessing forest condition can be applied with RS derived inputs for woodland assessment purposes over landscape-scale areas.

Table of contents

Copyright statement	ii
Abstract	iii
List of contents	iv
List of figures	xi
List of tables	xiv
List of formulae	xvii
Common abbreviations	xix
Preface/Acknowledgement	xxi

Main body of text

1	Introduction.....	1
1.1	<u>Rationale</u>	2
1.2	<u>The scope of this research project</u>	4
1.3	<u>Research aims and objectives</u>	5
1.4	<u>Structure of the thesis</u>	6
1.4.1	Part one: Context of the research.....	6
1.4.2	Part two: Dataset collection and analysis.....	6
1.4.3	Part three: Interpretation, synthesis and conclusion.....	7
2	Literature Review.....	8
2.1	<u>Definitions of Condition</u>	8
2.1.1	Ecological considerations in defining forest condition.....	9
2.1.2	Forestry definitions of condition.....	14
2.2	<u>Forest condition assessment</u>	17
2.2.1	Field based methods of forest condition assessment.....	21
2.2.1.1	<i>Fieldwork construction</i>	21
2.2.1.2	<i>Key-species sampling methods</i>	22
2.2.1.3	<i>Key-structures sampling methods</i>	22
2.2.1.4	<i>Additional metrics derived through instrumentation</i>	24
2.2.2	Woodland condition index construction.....	25
2.2.2.1	<i>Horizontal diversity indices</i>	26
2.2.2.2	<i>Vertical diversity indices</i>	26
2.2.2.3	<i>Complex diversity indices</i>	26
2.2.2.4	<i>Selection of indices to test</i>	28
2.2.3	Potential issues with the fieldwork approach.....	28
2.2.4	Assessment of forest condition within this project.....	29
2.2.4.1	<i>Woodland condition definition</i>	29
2.2.4.2	<i>Identification of required field metrics</i>	29
2.3	<u>Applications of Remote Sensing for Forest Condition Assessment</u>	32
2.3.1	Multispectral and hyperspectral Sensors.....	35
2.3.1.1	<i>Forest stand structure</i>	36
2.3.1.2	<i>Forest composition</i>	41
2.3.1.3	<i>Attribute summary</i>	43
2.3.2	Light Detection and Ranging (LiDAR)	44
2.3.2.1	<i>Types of airborne LiDAR systems</i>	44
2.3.2.2	<i>Applications of LiDAR Remote Sensing for ecology</i>	46
2.3.2.3	<i>Forest structure</i>	47
2.3.2.4	<i>Forest composition</i>	54

2.3.2.5	<i>Forest understorey and deadwood components</i>	57
2.3.2.6	<i>Attribute summary</i>	60
2.3.3	Fusion of optical and LiDAR data.....	60
2.3.3.1	<i>Forest structure</i>	61
2.3.3.2	<i>Forest composition</i>	62
2.3.3.3	<i>Summary of attributes</i>	64
2.4	<u>Conclusions</u>	64
3	Study Site and Datasets	67
3.1	<u>Study Site</u>	67
3.1.1	Field-plot site locations visited in 2010.....	71
3.1.2	Field-plot site locations visited in 2012.....	73
3.2	<u>Data Sources</u>	75
3.2.1	Forestry Commission inventory data.....	75
3.2.2	Field data capture.....	75
3.2.2.1	<i>Number of field plot samples</i>	76
3.2.2.2	<i>Field data plot size</i>	76
3.2.2.3	<i>Field data collection locations</i>	77
3.2.2.4	<i>Field data measurements</i>	80
3.2.3	Airborne remote sensing.....	84
3.2.3.1	<i>Archived airborne multispectral data</i>	87
3.2.3.2	<i>Airborne hyperspectral data</i>	87
3.2.3.3	<i>Airborne LIDAR Data (Discrete and Full-waveform)</i> ...	87
4	Methods	90
4.1	<u>Fieldwork Analysis</u>	94
4.1.1	Calculations applied to the field datasets.....	94
4.1.2	Hemispherical photography processing.....	94
4.1.3	Calculating compositional and structural indices.....	98
4.1.3.1	<i>Species composition indices</i>	98
4.1.3.2	<i>Structural indices</i>	99
4.1.3.3	<i>Combined indices</i>	101
4.1.3.4	<i>Score based indices</i>	102
4.2	<u>Hyperspectral Remote Sensing Processing</u>	103
4.2.1	Hyperspectral data pre-processing.....	103
4.2.1.1	<i>Cross-track correction</i>	104
4.2.1.2	<i>Geo-corrections</i>	104
4.2.1.3	<i>Atmospheric Correction</i>	105
4.2.1.4	<i>Between flightline radiometric difference correction</i> ...	106
4.2.1.5	<i>Image mosaicing</i>	107
4.2.1.6	<i>Quantification of horizontal accuracy</i>	107
4.2.1.7	<i>Image dimensionality reduction</i>	108
4.2.2	Spectral indices.....	110
4.2.3	Object-based tree species classification using hyperspectral data.....	111
4.2.3.1	<i>Development of the eCognition workflow</i>	113
4.2.3.2	<i>Creation of the forest mask</i>	114
4.2.3.3	<i>Individual tree crown delineation</i>	114
4.2.3.4	<i>Broad tree species classification</i>	116
4.2.3.5	<i>In-depth tree species classification</i>	117

	4.2.3.6	<i>Classification 30x30m summary metrics</i>	121
4.3		<u>Discrete return LiDAR processing</u>	122
	4.3.1	Filtering ground and non-ground points.....	122
	4.3.2	Ground separation and normalisation.....	125
	4.3.3	Extraction of DR LiDAR structure related metrics.....	125
	4.3.4	Individual tree crown detection.....	128
4.4		<u>Full-waveform LiDAR processing</u>	131
	4.4.1	Full-waveform Gaussian fitting.....	132
	4.4.2	Filtering and area-based metric extraction.....	135
	4.4.3	Individual tree detection.....	139
	4.4.4	Additional metrics for DR and FW data derived using the FW processing chains.....	139
4.5		<u>Data analysis methods</u>	140
	4.5.1	Field data analysis.....	140
	4.5.1.1	<i>Comparison of the 2010 and 2012 field populations</i>	140
	4.5.1.2	<i>Spearman's rho bivariate correlation</i>	141
	4.5.1.3	<i>Ordinary Least Squares (OLS) regression analysis</i>	142
	4.5.2	Hyperspectral analysis.....	144
	4.5.2.1	<i>OLS linear multiple regression</i>	144
	4.5.2.2	<i>Akaike's Information Criterion</i>	144
	4.5.3	Discrete-return and full-waveform LiDAR analysis.....	146
	4.5.3.1	<i>OLS linear multiple regression</i>	147
	4.5.3.2	<i>Akaike's Information Criterion</i>	147
	4.5.3.3	<i>Direct calculation of indexes from individual tree detection data</i>	147
	4.5.4	Combined Hyperspectral and LiDAR analysis.....	148
	4.5.4.1	<i>Defining combined variable subsets</i>	148
	4.5.4.2	<i>OLS linear and AICc regression</i>	148
4.6		<u>Validation using field data</u>	149
	4.6.1	Using remote sensing to predict field plot-level metrics.....	149
	4.6.2	Computation and validation of condition indices.....	150
4.7		<u>Mapping the results across the study site</u>	152
4.8		<u>Methods summary</u>	153
5		Fieldwork results	157
5.1		<u>Overview of fieldwork data</u>	157
	5.1.1	Statistical comparison of 2010 and 2012 fieldwork populations.....	172
5.2		<u>Relationships between field metrics</u>	174
	5.2.1	Bivariate correlation.....	174
	5.2.2	OLS multiple regression.....	175
5.3		<u>Condition index results for field plots</u>	178
5.4		<u>Summary of field plot data results</u>	185

6	Forest assessment using airborne hyperspectral data.....	187
6.1	<u>Object-based tree species classification.....</u>	187
6.1.1	Assessment against Forestry Commission stand inventory information.....	192
6.1.2	Comparison of species identified in the field versus validation plots.....	195
6.2	<u>Statistical models for estimating forest structural and compositional metrics from hyperspectral remote sensing.....</u>	198
6.2.1	Forest structure within the plot.....	198
6.2.2	Forest composition within the plot.....	202
6.2.3	Deadwood within the plot.....	204
6.2.4	Understorey metrics within the plot.....	205
6.2.5	Metrics required for condition index construction.....	208
6.3	<u>Validation of hyperspectral remote sensing derived model estimates against field data.....</u>	209
6.3.1	Direct comparisons between field measurements and hyperspectral remote sensing metrics.....	209
6.3.2	Validation of hyperspectral remote sensing derived statistical models against field data.....	213
6.4	<u>Summary of findings.....</u>	215
7	Forest assessment using LiDAR data.....	217
7.1	<u>Assessment of DR and FW LiDAR pre-processing data products.....</u>	217
7.1.1	Accuracy assessment of DTMs against GPS ground control....	217
7.1.2	Comparison of DR and FW small-footprint LiDAR outputs....	218
7.2	<u>Estimating field-level forest structural and compositional metrics from DR LiDAR data.....</u>	223
7.2.1	Forest structure within the plot.....	223
7.2.2	Forest composition within the plot.....	226
7.2.3	Deadwood within the plot.....	227
7.2.4	Understorey metrics within the plot.....	228
7.2.5	Metrics required for condition index construction.....	231
7.3	<u>Validation of DR LiDAR remote sensing derived model estimates against field data.....</u>	232
7.3.1	Direct comparisons between field measurements and remote sensing metrics.....	232
7.3.2	Validation of remote sensing derived statistical models against field data.....	235
7.4	<u>Estimating field-level forest structural and compositional metrics from FW LiDAR data.....</u>	237
7.4.1	Forest structure within the plot.....	238
7.4.2	Forest composition within the plot.....	241
7.4.3	Deadwood within the plot.....	242
7.4.4	Understorey metrics within the plot.....	243
7.4.5	Metrics required for condition index construction.....	246
7.5	<u>Validation of FW LiDAR remote sensing derived estimates of forest variables against field data.....</u>	247
7.5.1	Direct comparisons between field measurements and remote sensing metrics.....	247

7.5.2	Validation of remote sensing derived statistical models against field data.....	249
7.6	<u>Estimating or calculating metrics required for input into the complex stand diversity index (CSDI).....</u>	252
7.7	<u>Summary of findings.....</u>	255
8	Forest assessment through the use of a fusion between hyperspectral and LiDAR datasets.....	258
8.1	<u>Estimating field-level forest structural and compositional metrics from combined hyperspectral and LiDAR data.....</u>	258
8.1.1	Forest structure within the plot.....	258
8.1.2	Forest composition within the plot.....	262
8.1.3	Deadwood within the plot.....	264
8.1.4	Understorey metrics within the plot.....	265
8.1.5	Metrics required for condition index construction.....	267
8.2	<u>Validation of remote sensing derived field-level metrics via comparison with field measurements.....</u>	270
8.2.1	Hyperspectral and DR LiDAR derived statistical models.....	270
8.2.2	Hyperspectral and FW LiDAR derived statistical models.....	272
8.3	<u>A review of all regression models created from the five remote sensing datasets.....</u>	275
8.3.1	Forest structure within the plot.....	275
8.3.2	Forest composition within the plot.....	277
8.3.3	Deadwood within the plot.....	278
8.3.4	Understorey metrics within the plot.....	279
8.3.5	Metrics required for condition index construction.....	281
8.4	<u>A review and summary of model prediction accuracy of all remote sensing datasets.....</u>	282
8.5	<u>Summary of findings.....</u>	288
9	Mapping forest metrics and condition indices.....	290
9.1	<u>Mapping of forest metrics.....</u>	290
9.2	<u>Condition index mapping.....</u>	313
9.2.1	Vertical evenness (VE) index.....	315
9.2.2	The Clark-Evans aggregation index.....	317
9.2.3	The tree diameter differentiation index.....	319
9.2.4	The complexity index (HC).....	322
9.2.5	The complex stand diversity index (CSDI).....	324
9.2.6	The Target and accumulative scoring technique.....	328
9.2.7	Condition index overview.....	331
9.3	<u>A cross-comparison of condition index results.....</u>	332
9.3.1	Index categorisation.....	333
9.3.2	Assessment of index mapping results.....	339
9.4	<u>Summary of findings.....</u>	351

10	Discussion.....	354
10.1	<u>A review of hyperspectral derived outputs.....</u>	354
10.1.1	Species classification using hyperspectral data.....	355
10.1.2	Extraction of forest metrics relating to condition using hyperspectral data.....	358
10.1.2.1	<i>Overstorey.....</i>	359
10.1.2.2	<i>Understorey.....</i>	360
10.1.2.3	<i>Deadwood.....</i>	361
10.2	<u>A review of DR LiDAR derived outputs.....</u>	361
10.2.1	Automated individual tree crown extraction.....	361
10.2.2	Extraction of forest metrics relating to condition using DR LiDAR for direct measurements.....	362
10.2.3	Extraction of forest metrics relating to condition using DR LiDAR for Area-based (30x30m) metrics.....	364
10.2.3.1	<i>Structural metric comparison.....</i>	365
10.2.3.2	<i>Compositional metric comparison.....</i>	367
10.2.3.3	<i>Deadwood.....</i>	368
10.2.3.4	<i>Understorey.....</i>	368
10.3	<u>A review of FW LiDAR derived outputs.....</u>	369
10.3.1	A comparison between DR and FW LIDAR datasets.....	370
10.3.2	Extraction of forest metrics relating to condition using FW LiDAR for direct measurements.....	371
10.3.3	Extraction of forest metrics relating to condition using FW LiDAR for direct measurements area-based (30x30m) metrics.....	372
10.3.3.1	<i>Structural metrics.....</i>	373
10.3.3.2	<i>Compositional metrics.....</i>	374
10.3.3.3	<i>Deadwood.....</i>	375
10.3.3.4	<i>Understorey.....</i>	376
10.3.4	LiDAR sampling disparity.....	378
10.3.5	Concluding thoughts.....	379
10.4	<u>Exploration of the fusion of the airborne remote sensing datasets for assessing condition.....</u>	379
10.5	<u>A review of remote sensing derived condition indices.....</u>	390
10.5.1	Condition assessment methods tested.....	391
10.5.2	A comparison of the forest condition index methods.....	394
10.5.3	Assessing condition in the New Forest study site.....	396
10.6	<u>Implications and future work.....</u>	400
10.6.1	Usefulness in the wider field.....	400
10.6.1.1	<i>Ecology and habitat quality assessment.....</i>	401
10.6.1.2	<i>Habitat suitability modelling.....</i>	402
10.6.1.3	<i>Forestry.....</i>	402
10.6.1.4	<i>Further uses with minor modifications.....</i>	403
10.6.2	Potential improvements and future work.....	403
10.6.2.1	<i>ITC underestimation.....</i>	403
10.6.2.2	<i>ITC suppressed trees.....</i>	404
10.6.2.3	<i>Overstorey species classification.....</i>	405
10.6.2.4	<i>Understorey species classification.....</i>	406
10.6.2.5	<i>Hyperspectral derived metrics.....</i>	406
10.6.2.6	<i>Additional LiDAR metrics.....</i>	407
10.6.2.7	<i>FW LiDAR pre-processing.....</i>	408

10.6.2.8	<i>DR vs. FW LiDAR sampling disparity</i>	410
10.6.2.9	<i>Alternative methods of estimation</i>	410
10.6.2.10	<i>Field data enhancement</i>	411
10.6.2.11	<i>Alternative condition assessment methods</i>	412
10.6.2.12	<i>Transferability</i>	413
10.7	<u>Overview</u>	414
11	Conclusion	416
11.1	<u>Estimating forest metrics</u>	416
11.2	<u>Methods of assessing condition</u>	421
11.3	<u>Key findings</u>	424
11.4	<u>Key conclusions</u>	426
11.5	<u>Final thoughts</u>	427
12	Cited references	430
 <u>Appendices</u>		
Appendix A	Calculations applied to data collected in the field	450
Appendix B	Hyperspectral data	454
Appendix C	LiDAR data	460
Appendix D	R statistics script listing	479
Appendix E	Estimation of field plot level metrics and index construction	485
Appendix F	Comparison of mapping of forest metrics from different inputs	492

List of figures

<u>Figure 3.1</u> – Study site locations within the New Forest.....	69
<u>Figure 3.2</u> – Various photographs taken during fieldwork to exemplify the various forest structural types and environments surrounding the forest.....	70
<u>Figure 3.3</u> – Study site location within the New Forest (1).....	72
<u>Figure 3.4</u> – Study site location within the New Forest (2).....	74
<u>Figure 3.5</u> – Field plot locations.....	78
<u>Figure 3.6</u> – A comparison of field plot summary statistics between four sample plot sizes.....	79
<u>Figure 3.7</u> – Field plot design.....	80
<u>Figure 3.8</u> – Hemispherical photograph plot design.....	85
<u>Figure 3.9</u> – Example hemispherical photograph.....	85
<u>Figure 3.10</u> – The overlapping extents of the 2010 April and July airborne remote sensing acquisitions.....	86
<u>Figure 4.1</u> – A flowchart illustrating the processing and analysis steps for the field data.....	91
<u>Figure 4.2</u> – A flowchart illustrating the analysis steps for remote sensing datasets.....	92
<u>Figure 4.3</u> – Flowchart illustrating the selection of the best predictive models and the production of condition indices.....	93
<u>Figure 4.4</u> – The 5 key steps to analysing hemispherical photographs for forest canopies.....	96
<u>Figure 4.5</u> – Hemispherical photograph processing (1).....	97
<u>Figure 4.6</u> – Hemispherical photograph processing (2).....	97
<u>Figure 4.7</u> – An example of the ITC segments created through the crown-centroid detection approach, overlaid on the CHM.....	116
<u>Figure 4.8</u> – LiDAR point cloud pre-processing.....	124
<u>Figure 4.9</u> – LiDAR dataset problems.....	124
<u>Figure 4.10</u> – Normalised LiDAR data.....	125
<u>Figure 4.11</u> – ITC points and polygons generated through TIFFS, shown for conifer stands of plantation forest.....	130
<u>Figure 4.12</u> – SPDlib airborne waveform LiDAR processing chain. Modified from that presented in Bunting et al (in press).....	134
<u>Figure 4.13</u> – An example of one of the extracted waveforms.....	135
<u>Figure 4.14</u> – Illustrates the classification of the waveform-derived point cloud problems.....	138
<u>Figure 4.15</u> – Errors observed in initial FW metric extraction.....	139
<u>Figure 6.1</u> – Classification map part 1.....	188
<u>Figure 6.2</u> – Classification map part 2.....	189
<u>Figure 6.3</u> – Classification map part 3.....	190
<u>Figure 6.4</u> – A sample subset of the eCognition derived classification map.....	191
<u>Figure 6.5</u> – Examples of over segmentation resulting in poor tree crown delineation.....	192
<u>Figure 7.1</u> – An illustration of the locations of flight lines 11, 12 and 16 for 2010 leaf-on LiDAR data acquisitions.....	219
<u>Figure 7.2</u> – The locations of samples sites.....	220
<u>Figure 7.3</u> – Summary of height bin point counts for each of the sample areas in flight line 11.....	221
<u>Figure 7.4</u> – Summary of height bin point counts for each of the sample areas in flight line 12.....	221
<u>Figure 7.5</u> – Summary of height bin point counts for each of the sample areas in flight line 16.....	222

Figure 9.1 – The subset area for presenting predicted field-level metrics over wide areas.....	291
Figure 9.2 – The predicted number of tree stems per 30x30m.....	295
Figure 9.3 – The predicted number of tree stems of native species per 30x30m.....	295
Figure 9.4 – The predicted percentage of tree stems of native species per 30x30m..	296
Figure 9.5 – The predicted number of tree species per 30x30m.....	296
Figure 9.6 – The predicted average of tree stem nearest neighbour spacing (m) per 30x30m.....	297
Figure 9.7 – The predicted standard deviation of tree stem nearest neighbour spacing (m) per 30x30m.....	297
Figure 9.8 – The predicted average DBH (cm) of tree stems per 30x30m.....	298
Figure 9.9 – The predicted standard deviation of DBH (cm) of tree stems per 30x30m.....	298
Figure 9.10 – The predicted total basal area (m ²) of tree stems per 30x30m.....	299
Figure 9.11 – The predicted percentage of large trees (i.e. DBH 40>x<80cm) per 30x30m.....	299
Figure 9.12 – The predicted percentage of very large trees (i.e. DBH >80cm) per 30x30m.....	300
Figure 9.13 – The predicted percentage of large trees (i.e. DBH >40cm) per 30x30m.....	300
Figure 9.14 – The predicted SH index value (for species counts per stem) per 30x30m.....	301
Figure 9.15 – The predicted SI index value (for species counts per stem) per 30x30m.....	301
Figure 9.16 – The predicted SH index value (for species proportion defined by basal area) per 30x30m.....	302
Figure 9.17 – The predicted SI index value (for species proportion defined by basal area) per 30x30m.....	302
Figure 9.18 – The predicted average height of the first live branch (m) per 30x30m	303
Figure 9.19 – The predicted standard deviation of height of the first live branch (m) per 30x30m.....	303
Figure 9.20 – The predicted average horizontal crown area (m ²) per 30x30m.....	304
Figure 9.21 – The predicted standard deviation of horizontal crown area (m ²) per 30x30m.....	304
Figure 9.22 – The predicted total horizontal crown area (m ²) per 30x30m.....	305
Figure 9.23 – The predicted total of standing deadwood volume (m ³) per 30x30m...	305
Figure 9.24 – The predicted decay class of standing deadwood per 30x30m.....	306
Figure 9.25 – The predicted total of downed deadwood volume (m ³) per 30x30m...	306
Figure 9.26 – The predicted decay class of standing deadwood per 30x30m.....	307
Figure 9.27 – The predicted number of saplings per 30x30m.....	307
Figure 9.28 – The predicted number of saplings of native species per 30x30m.....	308
Figure 9.29 – The predicted number of saplings species per 30x30m.....	308
Figure 9.30 – The predicted number of seedlings per 30x30m.....	309
Figure 9.31 – The predicted number of seedlings of native species per 30x30m.....	309
Figure 9.32 – The predicted of seedlings species per 30x30m.....	310
Figure 9.33 – The predicted of SH index for native seedlings species per 30x30m...	310
Figure 9.34 – The predicted number of vascular species per 30x30m.....	311
Figure 9.35 – The predicted percentage of bare earth/soil per 30x30m.....	311
Figure 9.36 – The predicted difference of the sum of DBH (cm) differences per 30x30m.....	312
Figure 9.37 – The predicted Index of vertical separation per 30x30m.....	312

Figure 9.38 – FC compartment boundaries and primary species class overlaid upon a 1x1m CHM.....	314
Figure 9.39 – The VE condition index as calculated from ITC information for 30x30m areas.....	316
Figure 9.40 – The Clark-Evans aggregation index as calculated from ITC positional information for 30x30m areas.....	318
Figure 9.41 – The Diameter Differentiation index as calculated from statistically derived information from FW and hyperspectral datasets for 30x30m areas.....	321
Figure 9.42 – The HC index as calculated from statistically derived information for 30x30m areas.....	323
Figure 9.43 – The CSDI as calculated from statistically derived and ITC information from the remote sensing datasets for 30x30m areas.....	325
Figure 9.44 – The score index as calculated from statistically derived information from FW and hyperspectral datasets for 30x30m areas.....	331
Figure 9.45 – A subset of the study site area depicting the categorised VE index result.....	334
Figure 9.46 – A subset of the study site area depicting the categorised Clark-Evans aggregation index result.....	335
Figure 9.47 – A subset of the study site area depicting the categorised diameter differentiation index result.....	336
Figure 9.48 – A subset of the study site area depicting the categorised complexity index (HC) result.....	337
Figure 9.49 – A subset of the study site area depicting the categorised CSDI result	338
Figure 9.50 – A subset of the study site area depicting the categorised score index result.....	339
Figure 9.51 – The mean of all threecategorised condition index maps.....	347
Figure 9.52 – The mean of all three standardised condition index map.	348
Figure 9.53 – The standard deviation of the three standardised condition index maps.....	349
Figure 9.54 – The standard deviation of the three standardised condition index map. Annotations illustrate the areas with high standard deviations.....	350
Figure 10.1 – The score based index applied across the whole study site.....	397

List of tables

<u>Table 2.1</u> – FCS indicators for terrestrial habitats.....	20
<u>Table 2.2</u> – Example JNCC guidelines for SSSI assessment.....	20
<u>Table 2.3</u> – Indices used to quantify structural complexity of woodlands.....	27
<u>Table 2.4</u> – A listing of forest condition indicators assessed through field operations.....	30
<u>Table 2.5</u> – A list of the most common vegetation indices derived from spectral imagery.....	40
<u>Table 2.6</u> – Summary of tree canopy descriptor concepts.....	57
<u>Table 3.1</u> – Forest structure variables recorded in the field.....	81
<u>Table 3.2</u> – Ground vegetation and deadwood variables recorded in the field.....	81
<u>Table 3.3</u> – Disturbance and soil variables recorded in the field.....	82
<u>Table 3.4</u> – A summary of all remote sensing datasets currently acquired for the New Forest.....	86
<u>Table 4.1</u> – Summary of calculations applied to field recorded data.....	95
<u>Table 4.2</u> – Components of the CSDI of stand structural diversity.....	102
<u>Table 4.3</u> – List of ecological indicators analysed and their target values for the New Forest.....	103
<u>Table 4.4</u> – Methods of reducing image dimensionality in ENVI (version 4.7).....	109
<u>Table 4.5</u> – A comparison of Eigen values from leaf-on and leaf-off PCA and MNF transforms for the first 10 bands.....	110
<u>Table 4.6</u> – A summary list of successfully applied spectral indices using ENVI software.....	112
<u>Table 4.7</u> – Hyperspectral data derived image input list.....	114
<u>Table 4.8</u> – Class list and membership functions.....	119
<u>Table 4.9</u> – Extracted discrete return LiDAR metrics used in this study.....	127
<u>Table 4.10</u> – Mathematical descriptions of computed etrics.....	127
<u>Table 4.11</u> – All extracted ITC metrics for a single crown derived from TIFFS....	130
<u>Table 4.12</u> – ITC area summary values of ITC objects within each 30x30m grid cell.....	131
<u>Table 4.13</u> – Summary of extracted metrics from waveform LiDAR.....	138
<u>Table 4.14</u> – Testing the assumptions of a regression analysis.....	142
<u>Table 4.15</u> – Field measurements and remote sensing metrics for comparison.....	149
<u>Table 4.16</u> – Summary of inputs for full study site.....	153
<u>Table 5.1</u> – Summarised Descriptions for field plots visited in 2010.....	158
<u>Table 5.2</u> – Summarised Descriptions for field plots visited in 2012.....	161
<u>Table 5.3</u> – 2010 Field metrics summary.....	166
<u>Table 5.4</u> – 2012 Field metrics summary.....	167
<u>Table 5.5</u> – Summary values (mean, standard deviation and range) for the total number of field plots for each attribute in 2010 and 2012 data.....	168
<u>Table 5.6</u> – Summary values (mean, standard deviation and range) for the coniferous field plots for each attribute in 2010 and 2012 data.....	169
<u>Table 5.7</u> – Summary values (mean, standard deviation and range) for the deciduous field plots for each attribute in 2010 and 2012 data.....	170
<u>Table 5.8</u> – Summary values (mean, standard deviation and range) for the mixed field plots for each attribute in 2010 and 2012 data.....	171
<u>Table 5.9</u> – Statistical comparisons between 2010 and 2012 fieldwork populations.	173
<u>Table 5.10</u> – Spearman’s rho bivariate correlation matrix.....	176
<u>Table 5.11</u> – OLS regression model relationships between field metrics.....	177
<u>Table 5.12</u> – Indices derived from 2010 field data.....	180

<u>Table 5.13</u> – Indices derived from 2012 field data.....	181
<u>Table 5.14</u> – Score based condition assessment for 2010.....	183
<u>Table 5.15</u> – Score based condition assessment for 2012.....	184
<u>Table 6.1</u> – Comparison between 10 FC compartments and coincident remote sensing derived classified objects.....	194
<u>Table 6.2</u> – Comparison between fieldwork validation plots from 2012 and the remote sensing derived classification ITC map.....	196
<u>Table 6.3</u> – Forest structural metrics, R^2 and AICc weights.....	200
<u>Table 6.4</u> – Hyperspectral metric inputs for each forest structural metric regression model.....	201
<u>Table 6.5</u> – Forest composition metrics, R^2 and AICc weights.....	203
<u>Table 6.6</u> – Hyperspectral metric inputs for each forest compositional metric regression model.....	203
<u>Table 6.7</u> – Forest deadwood metrics, R^2 and AICc weights.....	204
<u>Table 6.8</u> – Hyperspectral metric inputs for each forest deadwood metric regression model.....	205
<u>Table 6.9</u> – Forest understorey metrics, R^2 , standard error and AICc weights.....	206
<u>Table 6.10</u> – Hyperspectral metric inputs for each forest understorey metric regression model.....	207
<u>Table 6.11</u> – Condition index construction metrics, R^2 , standard error and AICc weights.....	208
<u>Table 6.12</u> – Hyperspectral metric inputs for each forest condition index metric regression model.....	208
<u>Table 6.13</u> – The proportion of native species according to classification.....	210
<u>Table 6.14</u> – Comparison of diversity indices produced from remote sensing data and field data.....	211
<u>Table 6.15</u> – A summary of input metrics for the first component index of the CSDI relating to relative species abundance.....	212
<u>Table 6.16</u> – The OLS and AICc hyperspectral regression models. Part 1: overstorey composition.....	214
<u>Table 6.17</u> – The OLS and AICc hyperspectral regression models. Part 2: understorey composition.....	215
<u>Table 7.1</u> – RMSE and NRMSE of LiDAR DTMs in relation to ground GPS and total station measurements for all 95 values.....	217
<u>Table 7.2</u> – Summary of DR and FW number of recorded pulses and points by flight line subset.....	220
<u>Table 7.3</u> – Forest structural metrics, R^2 and AICc weights for DR LiDAR data...	224
<u>Table 7.4</u> – DR LiDAR metric inputs for each forest structural metric regression model.....	225
<u>Table 7.5</u> – Forest composition metrics, R^2 and AICc weights for DR LiDAR.....	226
<u>Table 7.6</u> – DR LiDAR metric inputs for each forest compositional metric regression model.....	227
<u>Table 7.7</u> – Forest Deadwood metrics, R^2 and AICc weights for DR LiDAR data...	228
<u>Table 7.8</u> – DR LiDAR metric inputs for each forest Deadwood metric regression model.....	228
<u>Table 7.9</u> – Forest understorey metrics, R^2 and AICc weights for DR LiDAR data	229
<u>Table 7.10</u> – DR LiDAR metric inputs for each forest understorey metric regression model.....	230
<u>Table 7.11</u> – Condition index construction metrics, R^2 and AICc weights for DR LiDAR.....	231
<u>Table 7.12</u> – DR LiDAR metric inputs for each forest condition index metric regression model.....	231

<u>Table 7.13</u> – Comparison of stem counts between field data and automated tree detection.....	233
<u>Table 7.14</u> – Comparison of forest structural metrics from DR LiDAR data with field data.....	234
<u>Table 7.15</u> – Canopy openness comparison of DR LiDAR data with field data.....	235
<u>Table 7.16</u> – The OLS and AICc DR LiDAR regression models. Part 1: overstorey composition.....	236
<u>Table 7.17</u> – The OLS and AICc DR LiDAR regression models. Part 2: understorey composition.....	237
<u>Table 7.18</u> – Forest structure metrics, R^2 and AICc weights for FW LiDAR data...	239
<u>Table 7.19</u> – FW LiDAR metric inputs for each forest structural metric regression model.....	240
<u>Table 7.20</u> – Forest composition metrics, R^2 and AICc weights for FW LiDAR data.....	241
<u>Table 7.21</u> – FW LiDAR metric inputs for each forest compositional metric regression model.....	242
<u>Table 7.22</u> – Forest deadwood metrics, R^2 and AICc weights for FW LiDAR data.	243
<u>Table 7.23</u> – FW LiDAR metric inputs for each forest deadwood metric regression model.....	243
<u>Table 7.24</u> – Forest understorey metrics, R^2 and AICc weights for FW LiDAR data.....	244
<u>Table 7.25</u> – FW LiDAR metric inputs for each forest understorey metric regression model.....	245
<u>Table 7.26</u> – Condition index construction metrics, R^2 and AICc weights.....	246
<u>Table 7.27</u> – FW LiDAR metric inputs for each forest condition index metric regression model.....	246
<u>Table 7.28</u> – Comparison of forest structural metrics from FW LiDAR data with field data.....	248
<u>Table 7.29</u> – Canopy Openness comparison.....	248
<u>Table 7.30</u> – The OLS and AICc FW LiDAR regression models. Part 1: overstorey composition.....	250
<u>Table 7.31</u> – The OLS and AICc FW LiDAR regression models. Part 2: understorey composition.....	251
<u>Table 7.32</u> – A summary of three minimum recorded NN distances for field measurements, LiDAR derived TIFFS ITC and corrected TIFFS ITC measurements.....	253
<u>Table 7.33</u> – A summary of three maximum recorded NN measurements for field measurements, LiDAR derived TIFFS ITC and corrected TIFFS ITC measurements.....	254
<u>Table 7.34</u> – A summary of the two minimum and two maximum tree crown horizontal diameters for field measurement and TIFFS ITC measurements.....	254
<u>Table 7.35</u> – A summary of approximated lowest crown base height value.....	255
<u>Table 8.1</u> – Forest structural metrics R^2 for hyperspectral and LiDAR combinations.....	259
<u>Table 8.2</u> – Hyperspectral and DR LiDAR metric inputs for each forest structural metric regression model.....	260
<u>Table 8.2</u> – Hyperspectral and FW LiDAR metric inputs for each forest structural metric regression model.....	261
<u>Table 8.4</u> – Forest composition metrics R^2 for hyperspectral and LiDAR combinations.....	262
<u>Table 8.5</u> – Hyperspectral and DR LiDAR metric inputs for each forest compositional metric regression model.....	263

Table 8.6 – Hyperspectral and FW LiDAR metric inputs for each forest compositional metric regression model.....	263
Table 8.7 – Forest deadwood metrics R^2 for hyperspectral and LiDAR combinations.....	264
Table 8.8 – Hyperspectral and DR LiDAR metric inputs for each forest deadwood metric regression model.....	265
Table 8.9 – Hyperspectral and FW LiDAR metric inputs for each forest deadwood metric regression model.....	265
Table 8.10 – Forest understorey metrics R^2 for hyperspectral and LiDAR combinations.....	266
Table 8.11 – Hyperspectral and DR LiDAR metric inputs for each forest understorey metric regression model.....	267
Table 8.12 – Hyperspectral and FW LiDAR metric inputs for each forest understorey metric regression model.....	268
Table 8.13 – Condition index construction metrics R^2 for hyperspectral and LiDAR combinations.....	269
Table 8.14 – Combined remote sensing metric inputs for each forest condition index metric regression model.....	269
Table 8.15 – The OLS and AICc hyperspectral and DR LiDAR regression models. Part 1: overstorey composition.....	271
Table 8.16 – The OLS and AICc for hyperspectral and DR LiDAR regression models. Part 2: understorey composition.....	272
Table 8.17 – The OLS and AICc hyperspectral and FW LiDAR regression models. Part 1: overstorey composition.....	273
Table 8.18 – The OLS and AICc for hyperspectral and FW LiDAR regression models. Part 2: understorey composition.....	274
Table 8.19 – Forest structural metric R^2 values for all datasets.....	276
Table 8.20 – Forest stand composition metrics R^2 values for all datasets.....	278
Table 8.21 – Deadwood within plot metrics R^2 values for all datasets.....	279
Table 8.22 – Regeneration and understorey metrics R^2 values for all datasets.....	280
Table 8.23 – Index Inputs metrics R^2 values for all datasets.....	281
Table 8.24 – The best regression models identified through interrogation of RMSE and NRMSE values. Part 1: overstorey composition.....	284
Table 8.25 – The best regression models identified through interrogation of RMSE and NRMSE values. Part 2: understorey composition.....	285
Table 8.26 – Summary of best statistical model choices.....	286
Table 8.27 – A breakdown of the regression models which best predicted each field-level metric.....	286
Table 8.28 – Regression model list for a combination of the ‘best’ FW, DR, hyperspectral or composite models.....	287
Table 9.1 – Summary of mapping results.....	292
Table 9.2 – Comparison of field vs. remote sensing derived VE index.....	315
Table 9.3 – Comparison of field vs. remote sensing derived Aggregation Index.....	317
Table 9.4 – Comparison of the field derived and remote sensing derived tree diameter differentiation index.....	320
Table 9.5 – Comparison of field vs. remote sensing derived complexity index.....	322
Table 9.6 – Comparison of field vs. remote sensing derived CSDI index.....	324
Table 9.7 – CSDI component index of tree species composition comparison between field and remote sensing derived indices.....	327
Table 9.8 – CSDI component Index spatial distribution comparison between field and remote sensing derived indices.....	327

<u>Table 9.9</u> – CSDI component index of crown differentiation comparison between field and remote sensing derived indices.....	328
<u>Table 9.10</u> – Comparison between the field and remote sensing derived score based index.....	329
<u>Table 9.11</u> – Decomposition of the combined ‘best’ Dataset score index.....	330
<u>Table 9.12</u> – The percentage of each of the categories for the six condition indices.	334
<u>Table 9.13</u> – Compartment composition for high mean values (> 4) within forest subset.....	343
<u>Table 9.14</u> – Compartment composition for low mean values (≤ 2) within forest subset.....	343
<u>Table 10.1</u> – The predictive regression models with the smallest RMSE.....	384

List of formulae

4.1 Shannon-Wiener index.....	98
4.2 Evenness index.....	99
4.3 Simpson index.....	99
4.4 Clark-Evans aggregation index.....	99
4.5 Clark-Evans aggregation index stochastic test statistic.....	100
4.6 Tree diameter differentiation index.....	100
4.7 Vertical Evenness index.....	101
4.8 The complexity index (HC).....	101
4.9 The complex stand diversity index (CSDI).....	101
4.10 Gaussian decomposition.....	133
4.11 The regression equation	143
4.12 Akaike's Information Criterionequation.....	144
4.13 Akaike's Information Criterionequation with correction for small sample sizes	145
4.14 Route mean square error (RMSE).....	150
4.15 Normalised route mean square error (NRMSE).....	150
6.1 Formula for determining Map Accuracy (MA).....	193
10.1 Index for return height diversity derived from LiDAR.....	408

Common abbreviations

AICc	Akaike's Information Criterion for reduced sample sizes
ARI	Anthocyanin reflectance index 2
ARSF	Airborne Research and Survey Facility
ARVI	Atmospherically resistant vegetation index
BA	Basal area
BNG	British national grid (coordinates)
CASI	Compact airborne spectrographic imager
CC	Canopy closure
CGV	Canopy geometric volume
CHM	Canopy height model
CMM	Canopy maxima model
CO	Canopy openness
Conif.	Coniferous forest type
CRI	Carotenoid reflectance index 2
CSDI	Complex stand diversity index
DBH	(Tree) diameter at breast height (i.e. 1.3m)
DDW	Downed deadwood
Decid.	Deciduous forest type
DEM	Digital elevation model
DR	Discrete-return
DTM	Digital terrain model
E	Evenness (index)
FC	Forestry commission
FLAASH	Fast Line-of-sight Atmospheric Analysis of Spectral Hyper-cubes (software)
FW	Full-waveform
GBH	Tree girth at breast height (1.3m)
GLA	Gap light analyser (software)
GME	Geospatial Modelling Environment (software)
GPS	Global positioning system
HC	Complexity index
HSCOI	Height scaled canopy openness index
ITC	Individual tree crown
JNCC	Joint nature conservation committee
LAI	Leaf area index
LAS	Log ASCII Standard (format file-type)
LiDAR	Light detection and ranging
MA	Map accuracy
MNF	Minimum noise fraction
MRENDVI	Modified red edge normalised difference vegetation index
MRESRI	Modified red edge simple ratio index
nDSM	Nomalised digital surface model
NDVI	Normalised Difference Vegetation Index
NERC	Natural Environment Research Council
NN	Nearest neighbour

NRMSE	Normalized Root Mean Square Error
OLS	Ordinary least squares (regression method)
OS	Ordnance survey (UK)
PCA	Principle component analysis
PRI	Photochemical reflectance index
R	Clark-Evans aggregation index
Radar	Radio detection and ranging
RENDVI	Red edge normalised difference vegetation index
RF	Random forest statistical algorithm
RGRI	Red green ratio index
RMSE	Root Mean Square Error
SH	Shannon-Wiener (index)
SI	Simpson (index)
SIPI	Structure insensitive pigment index
SPD	Sorted Pulse data (format file-type)
SPDlib	Sorted Pulse Software Library (software)
SRI	Simple ratio index
SSSI	Site of special scientific interest
StDev.	Standard deviation
TIFFS	Toolbox for LiDAR Data Filtering and Forest Studies (software)
TM	Tree diameter differentiation
VE	Vertical elevation (index)
VI	Vegetation index (from hyperspectral data)
VREI	Vogelman red edge index 1
WBI	Water band index

Acknowledgements

I am indebted to many people for making the time working on my PhD an unforgettable experience.

Foremost, I would like to offer my sincerest thanks to both my supervisors Dr Ross Hill and Dr Shelley Hinsely from Bournemouth University and the Centre for Ecology and Hydrology (CEH) for all their guidance, support and encouragement. I could not be prouder of my academic roots and hope that I can in turn pass on the research values that they have given to me.

In addition to my supervisors, I would also like to thank the numerous others whom have aided me over the course of my PhD project. The work undertaken could not have been performed without logistical and technical support. From Bournemouth University – Harry Manley and Karen Walmsley for their help with fieldwork equipment, and Shaun Bental for his technical assistance. From CEH – Stephen Freeman for his assistance with advice on statistical methods. I'd also like to thank the administration staff at Bournemouth University, most notably Louise Pearson, for their assistance.

I would like to extend special thanks to Dr Peter Bunting (from Aberystwyth University) for his patient assistance with his LiDAR remote sensing processing software, and to Emma Carolan from Plymouth Marine Laboratories (PML) for her invaluable assistance with Python code. I'd also like to thank Dr John Armston (University of Queensland, Australia) and Dr Martin Isenburg(USA) for responding to my queries about additional LiDAR software.

I would also like to thank the National Environment Research Council (NERC), in particular the airborne research and survey facility (ARSF) for providing my various remote sensing datasets, and the Field Spectroscopy Facility (FSF) for loaning me some very expensive equipment. In addition I am grateful to the Forestry Commission offices in the New Forest for their cooperation with my research, allowing me site access and keeping be informed about forestry activities.

I am also very grateful to all those who have assisted me to conduct my fieldwork in all weather conditions from blistering sun to torrential rain, and offer my apologies for the often ludicrous amounts of biting insects.

I would like to express a great deal of gratitude to my countless friends and family who have been supportive throughout completing my PhD. First I would like to thank my parents and my sister for their immense support, not least of which from their able fieldwork assistance. I would like to acknowledge the support of my friends, many of whom are also fellow PhD students in the school of applied sciences, for their advice, support and important distractions from the work to keep me relatively sane.

Chapter 1 – Introduction

This thesis discusses the use of combined remote sensing to estimate forest condition, taking a case study of the New Forest, Hampshire. The application of such technologies is an accurate means of assessing forest attributes and could provide an informative tool over larger areas.

Woodland covers about 10% of the UK's land area, but little of this is in a completely natural state. Those woodlands that resemble original forests are more highly prized in nature conservation terms. For the management of any woodland (whether conservation or production forests) an indication of woodland quality is required. Typically, the definition is based on a combination of the woodland's compositional and structural components, together with an assessment of disturbance, regeneration and naturalness. There is no single commonly used definition of 'quality' and it also varies between management objectives.

Currently woodland quality assessments are performed through fieldwork, where forest stand structure, composition, deadwood, regeneration, and ground vegetation are assessed. Fieldwork is hampered by cost and restricted spatial coverage, and is often neither objective nor repeatable. Recent developments have allowed the application of new sensors and methods in airborne remote sensing, offering an alternative or complement to fieldwork and providing a means of extrapolating metrics over large areas. Combining airborne hyperspectral and LiDAR data is expected to offer detailed and complementary information about the structure and composition of woodlands. This project proposes the use of a sensor-fusion approach to develop methods of estimating forest condition over wide areas.

This chapter provides an overview of the research project, beginning with the rationale for the research, and a general description of how the researcher has gone about this. The following section outlines the research aims and objectives, and finally the thesis structure.

1.1. Rationale

Forests and woodlands are important on a variety of levels. They are exploited as a resource for a wide range of human activities (Lucas et al., 2008a), as well as being important habitats for a variety of organisms and storing substantial amounts of above-ground carbon stocks (Horner et al., 2010). The concept of forest condition or quality is critical to the study of woodland ecology. Defining what ‘condition’ is in terms of the organism or community is complicated and problematic (Hinsley et al., 2002). For woodland these difficulties are compounded by the three-dimensional volume and complexity of the habitat (Hinsley et al., 2006). There are many differences in the definition of condition between research, conservation and commercial interests. Condition is often defined in terms of biodiversity, species richness or productivity dependent upon the management objective.

The relative condition value of certain sites is important in various conservation strategies. Forest managers cannot measure everything within a forest of interest, thus the choice of what to measure is critical. These measurements will form the indicators of condition, for example the monitoring of a certain type of species known to be sensitive to a certain process, or the fragmentation or isolation of an area (Noss, 1999; Vellend et al., 2007). The management of an area is often decided through landscape-scale monitoring, whereas ecological monitoring is often accomplished through extrapolating up from smaller scales based on field plot assessment (Newton et al., 2009b). With new conservation and monitoring initiatives being implemented through the European Union, such as Natura 2000 (JNCC, 2004), national scale reporting is required, thus the timeliness and accuracy of condition metrics becomes all the more important.

Temperate forests have been recognised as important components of the global carbon cycle. However, a comprehensive understanding of the overall spatial patterns of structural variation seen in these large landscapes is still largely incomplete (Anderson et al., 2008). Accurate structural information is required for a number of applications including forest management, environmental protection, fire behaviour analysis, and global warming and carbon management allocations (Chen et al., 2007).

Owing to the spatially heterogeneous nature of landscapes, ecological approaches can potentially be applied at a variety of scales, from an individual tree to the whole landscape, and to a wide range of different environments through the use of remote sensing techniques. Fine spatial resolution hyperspectral imagery and LiDAR data capture differing, yet complementary characteristics in the estimation of various vertical and horizontal forest structural metrics, and are often described as synergistic (Anderson et al., 2008; Coops et al., 2004).

The use of each of these systems addresses some of the limitations of the other. Information on individual tree species, health and spatial geometry of the tree crown can be provided by optical/spectral imagery, whereas LiDAR provides data concerning the tree height, 3D crown shape and texture or outline (Hyypä et al., 2004; Leckie et al., 2003b). Recent research projects, in particular for LiDAR, have focused upon the extraction of forest structural features (Hyypä et al., 2008; Lucas et al., 2008b; Maltamo et al., 2005). The use of airborne digital remote sensing technologies for forest ecology is an area of intense research, and there is little in the way of consistent guidance as to how best the data should be captured, processed and analysed.

A survey conducted by Newton et al. (2009b) suggested that landscape ecologists have been rather conservative in their use of remote sensing data sources. In addition, Aplin (2005) suggested that remote sensing specialists have focused on the technological issues as their principle concern, rather than ecological problems, which highlights a potential divide between the different research communities. There is therefore a need for closer integration and collaboration between these two disciplines. This could potentially offer the characterisation of horizontal and vertical components of forests of considerable value for understanding landscape structure and function. To date there has been little work undertaken to examine the potential for application of LiDAR and hyperspectral survey to characterise forested areas that exhibit a range of structural and species types, and management techniques. However, work on forests in Europe, America and Australia have demonstrated great promise in addressing this potential for the use of remote sensing to quantify forest structure and condition (Lucas et al., 2008b; Mason et al., 2003; Mutlu et al., 2008; Popescu et al., 2004).

1.2 The scope of this research project

This investigation will take place within the New Forest, Hampshire, in the south of the UK. This area is made up of a complex mosaic of forest structural types, including managed plantation, semi-ancient coniferous and deciduous woodland in close proximity to one another, and thus provides a suitable test site. This area provides many structural and compositional variables of interest. The New Forest is recognised as being of international importance for nature conservation. The specific study area has also been subject previously to a number of ecological remote sensing studies (Blackburn and Milton, 1996; Koukoulas and Blackburn, 2004, 2005a; Milton et al., 1994).

For the management of any woodland an indication of ‘quality’ is required. Recent developments have allowed the application of new sensors and methods in airborne remote sensing, offering an alternative to fieldwork and providing a means of intelligently estimating these metrics over large areas, rather than extrapolation from small samples. Additionally, remote sensing offers new possibilities when considering the national scale reporting requirements of existing conservation and monitoring initiatives.

There are many examples of research projects utilising remote sensing methods to extract metrics relating to forest composition or structure. There are very few studies existing that explicitly utilise remote sensing for the assessment and mapping of multiple woodland condition indicators. One such example is Simonson et al. (2013) who focused on combining spectral and LiDAR data to classify eleven land cover types, and using prior knowledge of the relationships between height, vegetation structure and species diversity was able to map three condition classes across a study area in Portugal. This project will adopt a similar methodology by combining hyperspectral and LiDAR data within the New Forest context.

The processing and usage of the relatively new technology of small-footprint full-waveform LiDAR is still a matter of intense research activity, and holds great potential for forestry applications (Adams et al., 2012). As such, many studies have been devoted to the technical aspects of the systems, and have routinely noted the potential advantages over conventional systems. There are a small, but growing number of studies documenting the output of such systems for forest applications, such as in Lindburg et al. (2012). None however have focused on deriving multiple forest related attributes. The study will address this by exploring the potential of these datasets.

In this investigation the researcher will explore the capabilities and methods of airborne remote sensing for the extraction of relevant metrics in order to assess condition for forests which can either be related directly to or be used to estimate those metrics recorded in the field. The complementarity of three remote sensing datasets: hyperspectral, discrete-return and full-waveform LiDAR, is analysed for the estimation of compositional and structural metrics. A methodology is demonstrated for the mapping of these forest condition indicators. The ultimate goal of the project is the creation of a suite of remote sensing derived indicators of condition and analysis methods designed to aid in the assessment of woodland condition over wide areas.

1.3 Research aims and objectives

The aim of this research is to assess the information content of combining airborne LiDAR and hyperspectral remote sensing systems, in order to develop and assess a suite of remotely sensed indicators of woodland condition which are objective and repeatable. This will be achieved through the following objectives:

1. To assess, through the recommendations of various advisory bodies and the scientific literature, the best forest condition indicators and the ideal assessment methods for application via field survey.
2. To review the utility of airborne remote sensing data for forest analysis and the understanding of ecological processes therein.
3. To assess the types of forest features that can be extracted from airborne LiDAR and optical-spectral remote sensing datasets.
4. To determine the best forest condition indicators that can be measured remotely, and which can be supported by fieldwork assessment.
5. To develop and test a workflow for processing airborne remote sensing data (hyperspectral and LiDAR) to identify relevant forest features.
6. To evaluate the results of the remote sensing to identify forest condition indicators (e.g. habitat and vegetation structure).

1.4 The structure of the thesis

The structure of the thesis is in the style of a report and is organised into eleven chapters, beginning with: (i) Introduction outlying the research aims and objectives, (ii) Literature Review, (iii) A description of Study Site and Datasets, (iv) Analysis Methods, (v-ix) Results, (x) Discussion and (xi) Conclusions. A summary of these chapters are outlined below.

1.4.1 Part one: context of the research

The first section provides an introduction to the research context and a methodological approach identified by the surrounding academic and public ‘grey’ literature, and discusses both its theoretical and practical applications. Chapter 1 introduces the topic, a basic rationale behind the research, and outlines the aims and objectives. Chapter 2 provides an in-depth review of the available surrounding literature. This chapter provides a more detailed background on forest management and the importance of woodland condition, in addition to an evaluation of current assessment methods. A review of potential alternative methods is provided for various remote sensing technologies and applications, examining their relevance for the assessment of forest condition from airborne platforms.

1.4.2 Part two: dataset collection and analysis

This section of the thesis begins with a description of the field study site, and of the data sources acquired for addressing the aims of the research project. Data sources from fieldwork, remote sensing and commercial organisations are summarised in Chapter 3. An overview of the two field campaigns is given, in addition to an overview of the forest metrics extracted. A summary of the remote sensing dataset specifications is also noted.

Chapter 4 summarises the methods used to process and analyse the datasets. It begins with the method of field data capture and the subsequent processing steps applied to derive various field metrics. The pre-processing and analysis methods applied to the remote sensing datasets follow. The methods of statistically analysing the remote sensing forest metrics and testing them against the field datasets are described. The end section summarises the process of mapping the condition indicators across the whole study site.

1.4.3 Part three: interpretation, synthesis and conclusion

The final section focuses on the presentation of results and discussion of the implications of this research project. Chapters 5 to 9 present the various results from the research project. Chapter 5 concerns the fieldwork results, site descriptions and computed condition index calculations in addition to reporting any relationships that exist between field variables. Chapter 6 presents the results for the hyperspectral processing, statistical analysis and validation. Chapter 7 presents the results for the LiDAR processing, statistical analysis and validation, in addition to comparing the differences between discrete-return and full-waveform outputs. Chapter 8 reports the results of fusion between both hyperspectral and LiDAR datasets, identification of the best overall models, statistical analysis and validation. Chapter 9 contains the results of the woodland condition indices derived from the remote sensing data and their validation, in addition to a number of site maps of condition.

Chapter 10 provides a discussion of the results and debates their position within the surrounding literature. The chapter also discusses the assessment of condition methods tested within this thesis. The chapter finishes with the identification of the original contribution that this research makes to its field and proposes several potential areas for future work. The final Chapter, 11, concludes with a consideration of the implications for future applications of woodland condition management by remote sensing approaches.

Chapter 2 – Literature review

This chapter contains a review of the surrounding academic and grey literature concerning the current definitions and techniques involved in assessing forest condition in order to complete objective one of this research project. To complete objective two a review was undertaken to identify examples from the surrounding remote sensing academic literature which could potentially lead to the detection or estimation of those attributes identified when answering objective one.

2.1 Definitions of Condition

The concept of condition or quality is critical to the study of ecology. Defining what condition is in terms of the organism or community is complicated and problematic (Hinsley et al., 2002). For woodland these difficulties are compounded by the three-dimensional structure and complexity of the habitat (Hinsley et al., 2006). There are many differences in the definition of condition between research, conservation and commercial interests. Condition is often defined in terms of biodiversity, species richness or productivity, amongst others, dependent upon the management objective. Nevertheless the relative condition value of certain sites is important in various conservation strategies.

An objective and realistic assessment of condition is a difficult concept. Different methods have yielded dissimilar results. If the analogy of ecosystem health or integrity is used, then an argument can be made that certain characteristics, such as the presence of large organisms or high species diversity, are desirable properties indicative of higher levels of health or integrity (Keddy and Drummond, 1996). The problem often relates to the existence of an operational target for woodland environments. For example when considering animal damage, where damage is classified as ‘a problem caused by an unwanted condition’, this is an anthropocentric concept used in relation to one or more specific species (Reimoser et al., 1999). To ascertain condition in an ecological system meaningfully requires that a specific aim – with a desired condition – be defined and compared with the current condition in order to determine whether habitat is improving or degrading. Noss (1990) recognised three primary attributes of ecosystems: composition, structure and function, these are defined as:

- *Composition* describes the parts of each biodiversity component in a given area (e.g. habitat types, species present, genetic populations within species).
- *Structure* refers to the physical characteristics supporting that composition (e.g. size of habitats, forest canopy structure, etc.).
- *Function* relates to the ecological and evolutionary processes affecting life within that structure (e.g. natural disturbances, predator-prey relationships, and species adaptation over time).

Ecosystem function can be considered as a temporal component and thus, requires time-series data in order to be quantified.

An analysis of current habitat states and impacts alone cannot lead to valid estimates of condition. It is only when a target has been explicitly stated that the actual status can be related to the probability of achieving that target (Reimoser et al., 1999). Socio-economic aspects and subjective valuations commonly play an important role in such assessments and subject to differing viewpoints, targets and thresholds, the estimate of what constitutes damage or benefit can be markedly different, even with identical levels of environmental impact.

2.1.1 Ecological considerations in defining forest condition

Recently biodiversity has become a major feature in conservation science, where it is often considered key in determining areas to spend resources. At present woodland biodiversity in the UK is in decline due to high rates of forest loss and degradation as a result of over-exploitation in some cases, neglect in others, and conversion of forest to other land uses (Newton et al., 2009b; Turner et al., 2003). Here ‘biodiversity’ is referred to as species and certain characteristics of species, primarily their distribution and number within a given area. In addition the use of biodiversity is meant more broadly to mean species assemblages and ecological communities – groups of interacting and interdependent species (Turner et al., 2003). Scale is an implied and key component of this definition (McElhinny et al., 2005).

Scientifically sound management requires frequent and spatially detailed assessments of species numbers and distributions (Turner et al., 2003). Such measurements can be prohibitively expensive to collect directly. Underlying this assessment of ‘condition’ is the assumption that certain key environmental parameters, which can be detected, will drive the distribution and abundance of species across landscapes and determine how they occupy

habitats (Turner et al., 2003). McElinny et al. (2005) state that the quantification of diversity should be made through the identification of the structural attributes for a forest stand. An effective and efficient biodiversity surrogate measure needs to be formed from an array of different structural variables, some of which were identified as: foliage arrangement; canopy cover; tree diameter; tree height; tree spacing; tree species; stand biomass; understorey vegetation and deadwood.

Most management programmes to conserve biodiversity focus upon the creation of protected areas. Conservation status is typically based on the assessment of landscape and ecosystem level features such as habitat loss, habitat fragmentation, the size and number of large habitat blocks, the degree of protection, and current potential threats (Noss, 1999). Loss of biodiversity and ecosystem integrity is often experienced in conjunction with a number of factors illustrating degenerative trends, for example: old forests have been replaced with younger forests and plantations; structurally complex stands have been replaced by simplified ones; large well connected patches have been replaced with smaller, more isolated patches; natural fires have been suppressed; many miles of road have been built in what were unbroken landscapes.

There are two general approaches to the measurement of biodiversity (Turner et al., 2003). The first is the direct measurement of individual organisms, species assemblages, or ecological communities. This depends upon the scale of measurement, for example surveying species occurrence or absence from a sample region. Alternatively, there is the indirect approach to monitoring biodiversity through the reliance on environmental parameters as proxies. Consider for example, many species are restricted to discrete habitats, such as woodland or grassland which can be identified at a broader scale. By combining information about the habitat requirements of species with maps of landcover, estimates of potential species ranges and patterns of species richness are available (Turner et al., 2003). Lindenmayer et al. (2006) state that spatial connectivity between habitats should be maintained, in addition to conservation of landscape heterogeneity and stand structural complexity, in order to better guide biodiversity conservation, which again engages the idea of the interplay between different scales.

Measuring progress or change of any kind requires the use of indicators. Typically forest managers will consider indicator species. Lindenmayer et al. (2006) state that many

indicators and criteria have been proposed to assess the sustainable management of forests; however their scientific validity remains uncertain. Because the effects of forest disturbance, such as logging, are often specific to particular species, sites, landscapes, regions and forest types, management through the use of indicator species, focal species, or threshold levels of vegetation cover are argued to be of limited generic value, controversial and difficult to select dependent upon the species. In many cases, attributes of a species' population, for example demographics, would be more useful in validating indicators rather than as indicators themselves (Noss, 1999).

Species loss is predominantly driven by habitat loss, and thus the overarching goal of conservation management must be to prevent this. The conservation of forest biodiversity will depend on the maintenance of habitat across a range of spatial scales. Newton et al. (2009a) outline several key principles which must be considered in the scope of monitoring for this objective. The first consideration is that of forest loss and fragmentation. Deforestation is typically accompanied by substantial forest fragmentation. The changes are associated with a decrease in percentage area of forest patches and an increase in isolation of those patches. It is necessary to consider the factor of connectivity. This is the linkage of habitats, communities and ecological processes at multiple spatial and temporal scales (Noss, 1999). Connectivity influences key biodiversity conservation processes, such as population persistence, recovery and disturbance, the exchange of individuals and genes in a population, and the occupancy of habitat patches (Lindenmayer et al., 2006; Newton et al., 2009a).

The characteristics of habitat edges are influenced by patterns of land use surrounding forest fragments and can have a major impact on biodiversity by affecting ecological processes such as dispersal, establishment, survival, and growth (Fuller, 2012; Newton et al., 2009a). Edge effects influence a variety of processes, including seed rain, seed germination, removal and predation, tree growth, animal movement and avian nest predation (Murcia, 1995). Newton et al. (2009a) identified that edge effects were influenced by human disturbance within the forest fragments, such as collection of firewood and livestock browsing. Indeed substantial forest biodiversity loss can occur due to human disturbance within the fragments themselves, through activities such as logging of timber, fuel wood cutting, livestock browsing, the development of infrastructure and fire setting.

Forest stand structural complexity embodies various stand attributes in addition to how they are spatially arranged within stands. Such factors contributing to stand structural complexity can include, for example, the diameter and age-class distributions of individual trees within the stand, in relation to those trees which survive or are removed. Additional factors are the spatial distribution of structural elements within the stand, the presence of large living trees, the presence of deadwood, the presence of gaps within the stand and age since formation. The vertical heterogeneity created from multiple or continuous canopy layers and horizontal heterogeneity including foliage density, canopy openness, and horizontal patchiness of profile types are of importance, for example these attributes have links to light penetration, and providing resources to various animal species (Lindenmayer et al., 2006; Noss, 1999). This structural complexity is critical for forest biodiversity conservation because it allows organisms to persist where they would otherwise be eliminated and facilitates a more rapid return of logged and regenerated stands to a suitable habitat condition for species which have been displaced (Lindenmayer et al., 2006).

Tree species richness can be influenced by management activities and disturbance events. Newton et al. (2009a) report that the total area, core area, edge length and proximity of forest fragments were all negatively associated with mean species richness of pioneer species, and positively associated with richness of forest interior species. Patch size appears to be the most important attribute influencing different measures of species composition. Forest fragmentation can also affect genetic variation within forest species, by influencing processes of gene flow, inbreeding and genetic drift.

Velland et al. (2007) indicate that past disturbance and management may cause community species composition across sites to become more or less homogenous. The author gives the example of the alpha and beta diversity of forest plants growing on former agricultural fields in contrast to older (ancient) forests in North-America and Europe. The presence of a number of ecological filters to colonisation may exist at any stage of the colonisation process for new sites. For example, recent forests may show reduced beta diversity if habitat specialists were less successful colonists than generalist species. Strong relationships between species isolation and species richness have been reported. Velland et al. (2007) present results suggesting that these filters, including dispersal, lead to reduced alpha diversity in recent forests and to homogenised species composition across the landscape. Human land uses would appear to make enduring changes in the patterns of biodiversity at multiple scales,

effectively decoupling species composition from environmental gradients (Devictor et al., 2008).

Noss (1999) also reports a number of impacts which may occur following human modifications to the environment and those related to climate change. Road construction for example has been correlated highly with disturbance levels, and in the case of road construction for the harvesting of timber, with habitat destruction. The invasion of exotic species and their dispersal by vehicles and equipment can alter community composition, cover and biomass levels. Increased air pollution, including low-level ozone, acid precipitation and particulates has impacts upon biomass increments and tree productivity due to changes in soil pH. and nutrient content. Direct damage to leaves and other tissues can occur. Global warming is also recognised as a modifier of forest condition status due to changes in temperature and moisture abundance. Forest systems may experience changes in biomass levels, productivity and species distributions as a result (Read et al., 2009).

It should be noted that the increasing use of forests for recreation, such as the activities of hiking, hunting, camping, etc. will alter the status of a forest site. Disturbance may take the form of displacing wildlife, footpath erosion, vegetation damage, changes in ground-level vegetation density and condition, and exotic species invasions (Noss, 1999).

It is recognised that ungulate herbivores can have a profound impact on the vegetation and soils within forests (Fuller and Gill 2001). Damage caused through twig browsing and bark peeling is an increasing problem in many European countries (Reimoser et al., 1999). Browsing and grazing from wild ungulates have always played a role in determining the structure and dynamics of natural ecological systems both in terms of their present day influence on the functioning of those ecological communities and as a powerful selection pressure in the original development of such systems (Putman, 1996).

Damage caused by large ungulates, such as deer, can occur at many levels, for example through trampling, feeding on the fruit or germinating seedlings, reducing the seed source and hampering natural regeneration, fraying, and through browsing or bark stripping of older trees that have survived the recruitment stage (Putman, 1996; Reimoser et al., 1999). Heavy grazing pressure can also result in dramatic changes in the composition and relative abundance of species of the woodland floor, even reducing diversity (Kirby, 2001). Grazing

does have positive effects via the maintenance of the heterogeneity of structure that many conservation managers seek to mimic: such as the opening of clearings, treading-in of seeds into the ground and their dispersal, and reducing canopy shade to permit the existence of shade-intolerant species (Putman, 1996; Reimoser et al., 1999). Further variability may be caused by the feeding of ungulates in one place and dunging in another, creating discontinuities in nutrient flows which may be detrimental or beneficial. It should be noted that wild herbivores, particularly deer, increase in population rapidly, often to the detriment of woodland habitats due to overgrazing (Quine et al., 2011). A long term study into Denny Wood in the New Forest, UK, a semi-ancient woodland, concluded that regeneration of tree species ceased after 1964, principally due to heavy grazing and browsing by deer and ponies (Mountford et al., 1999).

It can be argued that damage to forests is not solely dependent upon the number of ungulates present in an area, but rather a combination of environmental factors. Such factors include: the forest type, size of available area, availability of cover, habitat structure, and distance to preferred forages, etc. Therefore, more damage can be anticipated in areas where the ‘attractiveness’ of an area is high, but forage availability is low (Putman, 1996).

2.1.2 Forestry definitions of condition

Common changes in forests in Europe and North America over the past two centuries in addition to outright deforestation includes the loss of old forests, simplification of forest structure, decreasing size of forest patches, increasing isolation of patches, disruption of natural fire regimes, and increased road building, all of which have had a negative effect on condition (Noss, 1999). These trends can be reversed, or at least slowed, through better management practices.

As mentioned previously most conservation programmes have focused upon the creation of protected areas (Lindenmayer et al., 2006). Forest ecosystems are also increasingly expected to produce multiple goods and services, such as timber, biodiversity, water flows and sequestered carbon. While many of these are not mutually exclusive, they cannot be simultaneously maximised so management compromise is inevitable and this involves a great deal of uncertainty (Horner et al., 2010). How management will in turn affect the condition of the forest system is a question of prime importance.

Rombouts et al. (2008) argues a definition of site condition where condition equates to productivity. Commercial forestry organisations outline the need for sustainable forestry (Treitz and Howarth, 1999), where the ability of the forest to sustain itself ecologically and provide societal needs is what defines a healthy forest. Maintaining the balance between forest sustainability and the production of goods and services is the challenge for managers of forest sites. These components are intrinsically linked. Many organisations have developed suites of indicators to monitor forest and environmental condition. The UK's Forestry Commission has a suite of ten broad indicator groups including air pollutants, water yield, river habitat quality, and damage through wind and fire. However, the importance of each of the indicator groups is not identical; importance is given to those representing the condition of the trees themselves, as well as a reflection of the health of the wider forest environment (Forestry Commission, 2011).

Initial site conditions and management practices have profound impacts on the final quality of timber produced. Vellend et al. (2007) explore the impact of past management practices from the clearance of forests for agriculture to the recovery and re-plantation of forests, and its impacts upon the biodiversity patterns for a number of forests in Europe and the USA. The findings from this research identify the importance of a number of factors such as soil properties and the dispersal of seeds within the environment which can impact on beta diversity. This is in part because dispersal plays such an integral part in forest (re)colonisation.

Forest structure is manipulated in two fundamental ways in plantation forests, the first being the initial planting density, and the second being the subsequent thinning operations (Florence 2004). Horner et al. (2009) report that lower density planting produced stands with higher mean and maximum stem diameters than higher density stands in an Australian forest in proximity to the Murray River. Stand density is also manipulated routinely by thinning in many forests to accelerate tree growth, minimise fire risk and restore forest structural complexity (Horner et al., 2010). When stands are thinned before annual biomass production has peaked, there is an opportunity to accelerate growth in the retained trees, increasing crown expansion, diameter growth and the rate at which the stand collectively reaches maturity. It should be noted that the presence of larger trees is associated with more hollows and a wider range of hollow sizes than stands with smaller trees. Several studies have shown

links between faunal diversity and the abundance of large tree hollows (Lada et al., 2007; Newton, 1994).

The long term effects of thinning on habitat quality and standing above-ground carbon stocks of forests is poorly understood (Horner et al., 2010). There have been a number of concerns about the characteristics and properties of the timber to be harvested in the UK. According to MacDonald and Hubert (2002), the general trend is for increasing knot sizes, reduced wood densities, increased juvenile core size and higher percentages of compression wood. These problems are associated with a progressive widening of initial spacing over the years and the effects of wind exposure.

Reimoser et al. (1999) stated that an objective and realistic assessment of damage is difficult, particularly as regards to browsing on natural regeneration. The problem often relates to the absence of an operational target. Damage being an anthropocentric concept, it is necessary to define a wanted and a current condition in order to determine whether ‘damage’ or merely an impact or disturbance has occurred. Damage to forests, as opposed to impacts, embraces concepts such as loss or reduction of increment, economic value, ecological stability, diversity, etc.

Target values for regeneration may for example be set for different stand types, such as setting the lowest acceptable number of tree stems or the distribution of tree species (Reimoser et al., 1999). In the establishment of a damage threshold it is essential to distinguish whether regeneration targets have been set, for example in terms of forest inventory requirements (i.e. the optimisation of forest income) or in social terms (i.e. the sustained protective forest function or landscape design). The effects of ungulates can best be judged by comparing the lowest acceptable regeneration with and without ungulate impact, e.g. by comparing inside and outside of an ungulate proof fence.

2.2 Forest condition assessment

A primary requirement of forest management is to assess the status, condition, and conservation value of the forest in question, at whatever spatial scale, relative to other forests (Noss, 1999). Knowledge of what kind of forest occupied the same regions in history, prior to intensive human activity, is important for the tracking of changes which have occurred and aiding in the prediction of future events (Noss, 1999). Thus monitoring of forest environments is a long-term process, often with no immediate results. Being fully aware of forest status, conditions, trends and conservation value is an enormous task. The consensus to make the process of assessment and monitoring less overwhelming and more practical to forest managers is to determine measurable indicators that correspond to the elements of forest condition, biodiversity, health, and sustainability that society finds valuable (Noss, 1999). Managers cannot measure everything of potential interest, thus the choice of what to measure is critical.

For management of forest activities the operational units are often defined at the stand-level. Stands are often the basic survey units for forest inventory. A stand is defined as a contiguous group of trees sufficiently uniform in species composition, arrangement of age classes, site quality and condition, to be a distinguishable unit (Pascual et al., 2008). At the scale of a stand, condition indicators are usually placed into one of two categories: those based on the identification of key species, and those based upon the identification of key structures (McElhinny et al., 2005). It should be noted that some approaches utilise a combination of both key species and key structures as indicators.

Measuring progress or change of any kind requires the use of indicators, for example the identification of an indicator species. Selection of the best indicator species, or species population, for the required purposes is difficult and controversial (Noss, 1999). Nevertheless, it often makes sense to monitor species populations directly. Key species are often defined as those most influenced by threatening processes, for example those taxon or taxa most limited by dispersal abilities (Lindenmayer et al., 2006).

Suites of key species are often used for indicators when managing a landscape, each of which is thought to be sensitive to a particular threatening process. There are serious flaws however; suites of key species are presumed to act collectively as a surrogate for other elements of the

biota, but a landscape managed for a given set of key species may not meet the requirements of the remaining biota (Lindenmayer et al., 2006). Key species indicators may also be hindered by the lack of scientific evidence to guide the reliable selection of indicators. Indicators based upon species presence or absence data in isolation to structural variables have had limited success because robust relationships between potential indicator species, or groups of species, and habitat condition or even biodiversity have not been established (Lindenmayer et al., 2006; McElhinny et al., 2005).

Key structures, or structural diversity, refers to the physical organisation or pattern of a system, including the spatial patchwork of different physical conditions in a landscape, habitat mosaics, species assemblages of different plant and animal communities, and genetic composition of subpopulations (Stokland et al., 2003). Indicators based upon key structures have been generating considerable interest. The rationale for this approach is that ecosystems containing a variety of structural components are considered likely to have a variety of resources and species that utilise these resources (McElhinny et al., 2005). Thus, there is often a positive correlation between elements of condition and the measures of variety and/or complexity of structural components within an ecosystem. From an ecological point of view, forest structure at the stand-level is of special interest when considering disturbance dynamics, growth stages and wildlife habitats (biodiversity issues) (Maltamo et al., 2005). Forest structures vary within and between vegetation zones around the world and are greatly influenced by silvicultural operations. They represent an indirect approach, as they show how the woodland may be composed, but give no information on whether the available resources are utilised (Christensen et al., 2004).

Tree and shrub species composition and structure are often cited as key indicators of the quality of woodland habitat for a wide range of organisms (Hill and Thomson, 2005). Additional characteristics such as woodland size, age, spatial arrangement, fragmentation, past disturbance, past management regimes, species competition and climate all contribute. Forest stand structure includes both vertical and horizontal layers. The former, concerns the number of tree layers and understorey vegetation, while the latter covers the spatial pattern of ground flora, trees and gaps. Additionally species richness is included (Maltamo et al., 2005; Pascual et al., 2008).

Simplification of species composition is also a concern, where structurally complex native forests are converted to simplified secondary stands or plantations over time. Moreover, fragmentation and isolation effects are of importance where remaining tracts of native forest are small and separated by terrain that is hostile to many species and pose barriers to movement (Noss, 1999; Vellend et al., 2007). The act of disturbance by other organisms should also be considered (Zerger, et al., 2009).

Stand structure is commonly defined in terms of two components in the ecological literature – stand structural attributes and stand structural complexity. The former can include measures of:

- *Abundance*: e.g. the density of tree volume and deadwood (Horner et al., 2010; Kim et al., 2009b).
- *Relative abundance*: e.g. foliage height density, DBH diversity, basal area of species groups (McElhinny et al., 2005).
- *Richness*: e.g. overstorey species richness, shrub species richness, etc. (Cantarello and Newton 2008).
- *Size Variation*: e.g. standard deviation of DBH.
- *Spatial Variation*: e.g. coefficient of variation of distance to nearest neighbour (Hill et al., 2010; McElhinny et al., 2005).

Attributes that quantify variation are particularly important because these can also describe habitat heterogeneity at the stand scale. Stand structural complexity is essentially a measure of the number of different structural attributes present and their relative abundance. McElhinny et al. (2005) state that structural complexity is used in preference to structural diversity because the latter term is considered ambiguous. This reflects the work of a number of authors in which a diversity measure, such as the Shannon-Weiner Index, has been used to quantify a single attribute, such as variation in stem diameter. The quantity is then deemed to be a measure of structural diversity and to be indicative of biological diversity. However, all that has been quantified is one of many possible attributes.

The overall richness of wildlife communities can be highly dependent on the three-dimensional spatial pattern of vegetation, especially in systems where biomass accumulation is significant (Lefsky et al., 2002). Individual bird species, in particular, have associations with the three-dimensional features within forests (Fuller, 2012; MacArthur and MacArthur 1961). In addition, other functional aspects of forests, such as productivity, may be related to forest structure.

As mentioned previously, Cantarello and Newton (2008) utilised four main groups of quality indicators in their field work assessment, these are summarised in Table 2.1. In contrast Table 2.2 summarises an example from the JNCC guidelines for SSSI assessment.

Table 2.1 – Favourable conservation status indicators for terrestrial habitats (Cantarello and Newton, 2008)

Key Factor:	Indicators:
<i>Forest Structure</i>	Number of tree species; basal area; mean stem diameter; tree heights; number of saplings; and indices of native to foreign tree species.
<i>Dead Wood</i>	Volume of downed deadwood; volume of standing deadwood (snag); and proportion of decay class.
<i>Tree Regeneration</i>	Number of total seedlings; and indices of native to foreign species.
<i>Ground Vegetation</i>	The number of ground vegetation species.

Table 2.2 – Example JNCC guidelines for SSSI assessment (JNCC, 2004)

Attribute:	Indicator Target:
<i>Area</i>	The existence of ancient or 19 th Century broadleaved woodland and its extent.
<i>Natural Processes and Structural Development</i>	No evidence of recent felling of native trees; little ground disturbance; no evidence of recent planting; no evidence of recent drainage/ditch maintenance; no evidence of forestry activities; no evidence of essential safety work (e.g. felling); canopy cover over 30-90%.
<i>Regeneration Potential</i>	At least 1 native sapling within 30 minutes of walk start; oak and beech
<i>Composition</i>	Little or no non-native species in the canopy or shrub layers; occasional (non-dominant) holly thickets (<50% ground cover).
<i>Characteristic Features</i>	<55% trees >80 cm DBH showing severe stress or death attributable to disease or pollution. Damage level inflicted by wildlife; <10% soil surface poached or trampled; <50% of vegetation more than 10cm high and <50% more than 40cm high; <10 of vegetation heavily modified, improved or exhibiting disturbed communities attributable to recreation activities.

2.2.1 Field based methods of forest condition assessment

Current assessment of woodland condition is primarily conducted via fieldwork. Because there is no consensus concerning which indicator to use in a given situation for monitoring of forest woodland condition, fieldwork campaigns can vary extensively between projects. Many countries and organisations utilise a set of guidelines for the capture of information, where methods can vary as long as the result is of a standard which can be compared relatively to other sites. One such example in the UK is that of JNCC Common Standards Monitoring (CSM) guidance (JNCC, 2004). The vast majority of woodlands within the UK are managed in some form. It should be noted that the nature of management surrounding these forests can also influence the objectives by which the fieldworkers judge the condition of a particular site (JNCC, 2004).

2.2.1.1. Fieldwork construction

Quantification of forest structure in the field typically involves multiple measurements of features such as stem sizes and densities, estimation of plant species composition, and light penetration through the canopy. Fieldwork typically takes the form of establishing a number of field plots throughout the study area where there are no obvious forest gradients (Cantarello and Newton, 2008). A sample plot size can vary significantly between projects, usually corresponding to a rectangular or circular area of 1ha or smaller, from within the forest stand (Kim et al., 2009b; Zimble et al., 2003). The number of sample plots varies, typically consisting of greater than ten permanent sample sites, covering all major structural types or features of interest (Cantarello and Newton 2008; Kim et al., 2009b; Pesonen et al., 2008). For extended research projects, mobile sample sites have been instituted (Pesonen et al., 2008). Selection of plot location can be accomplished through a variety of different sampling designs, or focussed around a feature of interest, such as the location of bird nesting sites (Clawges et al., 2008). Sample plot locations are referenced to map coordinates, commonly through Global Position System (GPS) devices (Cantarello and Newton, 2008; Kim et al., 2009). Within each plot the condition indicator variables can be collected, some of which are collected within a smaller sub-plot or transect within the main plot due to the potentially huge amount of ground vegetation encountered, for example in the number of ground flora species.

Alternative fieldwork methods exist, such as the point-transect-based method, which is appropriate for the estimation of population density of static objects, such as trees, in inaccessible areas or when time is a limiting factor (Cantarello and Newton, 2008). Another alternative is that of simple visual assessment; this method involves the surveyor completing a questionnaire during a structured walk. Such a questionnaire can be designed, for example, to implement JNCC CSM guidance for woodland habitats (JNCC, 2004).

2.2.1.2 Key-species sampling methods

When considering using key species as condition indicators, measurement must be accomplished in a different manner. The most common method is to infer quality retrospectively from bird breeding or territorial occupancy (Hinsley et al., 2009a; Hinsley et al., 2002). This can be accomplished through producing a census of the species population within the study area. For example territorial bird species were trapped and colour-ringed over a period of six years by Broughton et al. (2006) in order to determine compositional and structural characteristics of viable territories for the birds. Measurements were then taken from the birds before release, such as weight, wingspan, age and gender. Additionally assessments such as locations where singing or territorial disputes were observed were also noted. Hinsley et al. (2006; 2009) utilised measurements from birds breeding in nest-boxes located throughout Monks Wood, UK, where mean chick body mass for each brood was used as an indicator of foraging conditions within the territory. Visual or auditory assessment has also been used to make counts of bird populations within forests, recording species and sex across multiple site visits to estimate population densities, e.g. Clawges et al. (2008).

2.2.1.3 Key-structures sampling methods

Traditional methods of deriving information on forest stands utilise sampling designs with transects, or alternatively random or systematically selected plots so that the final stand parameters can be derived based on statistical extrapolation (Wang et al., 2004). Even when utilising remote sensing, all forest inventories require the use of a fieldwork component. Typically fieldwork campaigns are designed to capture a representative range of characteristics for the various trees within the sample area. Two of the major characteristics are stem Diameter at Breast Height (DBH) and height of the main section of trunk. Additional measures such as crown size, vegetative strata, tree age and species type can be recorded (Thomas et al., 2008). Other environmental factors may also be recorded depending on the purpose of the survey, such as soil type and climate (Campbell and Wynne 2011).

These surveys are designed to ensure that all major allometric tree species associations are captured for the given environment (Rosette et al., 2008; Thomas et al., 2008).

Various measures such as stem DBH and height can be taken in the field with the use of clinometers and distance tape (Coops et al., 2004), and form the basis of calculating stand metrics such as mean height, basal area and Lorey's average height (Hall et al., 2005). Recently more technologically advanced techniques have been applied, including GPS receivers and laser rangefinders to determine the locations of trees and their heights (Anderson et al., 2008). However, difficulties arise in determining tree height, as the apex can often be obscured and even merged with another tree crown (Boudreau et al., 2008; Lucas et al., 2008b). GPS receivers suffer positional inaccuracy within forests unless used with a large antenna taller than the surrounding trees. The accessibility of the site is also an important concern (Campbell and Wynne 2011). Site knowledge should also not be underestimated, as often what is documented is related to the knowledge of the surveyor.

The field measures can then be used to develop allometric calculations to estimate characteristics for other similar trees in the stand or ecosystem. For example canopy structure, which is essentially the sum of the sizes, shapes, and relative placements of the tree crowns in a forest stand, is central to many aspects of ecology (Purves et al., 2007). These equations are typically nonlinear and relate measures of tree size (such as DBH, tree height, canopy structure and crown area) to the dry weight of the above-ground (leaves, branches, trunks) and/or below-ground (roots) components of biomass or the sum of these in total (Lucas et al., 2008b; Purves et al., 2007). These allometric relationships are determined typically for a particular species by destructively harvesting trees across the size range, as observed in the field or derived from existing forest inventory data through statistical analysis (Lucas et al., 2008b; Popescu et al., 2004). Tree height may be used as an input to allometric equations, most however utilise DBH, which is typically measured 130cm above-ground level. Problems arise with multi-stemmed individuals. In many cases DBH can be estimated by its species-specific relationship with height (Lucas et al., 2008b). Crown area is sometimes used as an input for allometric equations, particularly where individual trees support a large number of stems, although obtaining reliable ground-based measures for crown areas is often problematic (Lucas et al., 2008b).

Allometric equations are typically site-specific. Different environmental factors can be more or less conducive to tree growth and biomass accumulation (Drake et al., 2003). Problems can also arise when considering differences with the species and when allometry is applied outside of the range of the calibration data causing incorrect estimation of tree characteristics. The major practical constraints are collecting sufficient data from which to determine these allometric relationships (Koukoulas and Blackburn 2005a).

2.2.1.4 Additional metrics derived through instrumentation

Forest canopy structure represents the complex spatial arrangement of foliage, branches, and boles of the trees; it has been a significant focus of research because of its influence on a wide range of biophysical and ecological properties (Frazer et al., 2001). The relationship of canopy structure and the temporal and spatial distribution of the incident understorey light have been involved in the study of natural disturbance, forest succession, timber harvesting and silvicultural prescriptions on the survival, pattern, and diversity of understorey plants and trees. Numerous ground-based optical tools and techniques have been developed to measure attributes of forest canopy structure. One optical method that has received increased attention is hemispherical (fisheye) canopy photography, because of its unique ability to permanently record the spatial properties of all visible canopy elements, and to rapidly predict seasonal flux of direct and diffuse light through discrete openings in the canopy (Frazer et al., 2001; Rautiainen et al., 2007).

One of the most crucial and most uncertain steps in the estimation of canopy structure metrics from hemispherical photography is the thresholding stage, in which the image is segmented into foreground (canopy) and background (sky) (Cescatti, 2007; Frazer et al., 2001). Owing to the high number of factors affecting image grey levels within these photographs, for example lens vignetting, gamma correction, heterogeneity in sky irradiance, etc., the estimation of canopy variables is rather uncertain with current thresholding techniques. Both manual and automatic segmentation techniques are highly dependent upon camera exposure and operator experience, and therefore can produce rather subjective and non-repeatable results. Cescatti (2007) proposes a solution for this issue, where the standard in-camera logarithmic conversion (gamma correction) is replaced with a linear conversion of the sensor analog signal. Thus the ability to acquire hemispherical photography with digital numbers proportional to radiance opens new possibilities in the analysis of canopy structure and microclimate, with important feedbacks in ecology and remote sensing.

Alternatively, devices such as the LAI-2000 Plant Canopy Analyser (Li-Cor Inc. Nebraska, USA) instrument can be used to estimate vegetative biophysical variables. The LAI-2000 requires measurements to be made from above and below the forest canopy, using a fisheye sensor to measure light interception at five angles, from which LAI can be estimated using a model of radiative transfer within the vegetation canopy. Other biophysical variables can be extracted such as foliage clumping and solar radiation regimes, and are reviewed elsewhere (Cescatti, 2007; Frazer et al., 2001; Rautiainen et al., 2010; Rautiainen et al., 2007).

2.2.2 Woodland condition index construction

Woodland structural diversity or complexity is defined as the composition of biotic and abiotic components in woodland ecosystems (Lexer et al., 2000), the specific arrangement of components within that system (Gadow, 1999), or as their positioning and combination (Heupler, 1982 as cited in Lübbert, 1999). The structure can also be characterised horizontally, i.e. the spatial distribution of the individuals (Zenner, 1999), and vertically in their height differentiation. An alternative is to define the structure as spatial distribution, mixture and differentiation of trees within a woodland ecosystem.

A number of methods exist to describe the structure and composition of woodlands utilising fieldwork captured data as input. The ‘classical’ methods of conveying a qualitative description either verbally or graphically may not be sufficient to reveal subtle differences (Kint et al., 2000). Therefore a number of quantitative methods have been proposed which should overcome this issue, as described in McElhinny et al. (2005), Neumann and Starlinger (2001) and Merganič et al. (2012).

A stand-level quantitative index of structural complexity is a mathematical construct which summarises the effects of two or more structural attributes in a single index (numeric) value (McElhinny et al., 2005). It is believed that as a summary value for a larger group of structural metrics, such an index could function as a reliable indicator of stand-level biodiversity and provide a means to rank stands in terms of their potential contribution to biodiversity (Neumann and Starlinger, 2001).

It is worthy of note that some authors have erroneously used a diversity measure, such as Shannon-Weiner index to quantify a single attribute and have then termed this attribute an index of structural complexity, when they have quantified only one of many possible attributes (McElhinny et al., 2005).

There is little consensus in the literature as to how to construct a complexity index, there is also a tendency for researchers to tailor their indices to suit their immediate study needs. As a result the surrounding literature contains a variety of different indices with no single index preferred over the others. The most prominent indices are summarised in Table 2.3; a selection of which are described in further detail below.

2.2.2.1 Horizontal diversity indices

The indices characterising woodland structure usually compare a hypothetical spatial distribution with the real situation (Neumann and Starlinger, 2001). Probably the most well-known index is the aggregation index (Clark and Evans, 1954), which describes the horizontal tree distribution pattern. It is a measure of the degree to which a forest stand deviates from a Poisson forest, where all individuals are distributed randomly (Tomppo, 1986). It is defined as the ratio of the observed mean distance to the expected mean distance if individuals were randomly distributed.

2.2.2.2 Vertical diversity indices

While there are many indices which measure horizontal structure, there are only a few for vertical structure. One such example is the index of vertical evenness, as put forward in Neumann and Starlinger (2001). The index utilises a Shannon index and stratification of individual tree heights into height layers, which characterise the vertical distribution of crown-horizontal coverage within a stand.

2.2.2.3 Complex diversity indices

Complex indices combine several biodiversity components in one measure. These indices are usually based upon an additive approach – where the final value is obtained as a sum of the values of individual biodiversity components. One such example of assessing biodiversity using a simple scoring method was outlined in Van Dem Meerschaut and Vanderkhove (1998). The authors developed a stand-scale forest biodiversity index based upon available data from forest inventory. The index combines four main aspects of forest ecosystem

biodiversity: forest structure, woody and herbaceous layer composition, and deadwood. Each aspect contains a set of indicators, for example forest structure is defined by canopy closure, stand age, number of stories, and spatial tree species mixture. The indicators are given a score determined on the basis of common agreement of their value. The sum of the eighteen scores (max 100) is the index value.

Table 2.3 – Indices used to quantify structural complexity of woodlands

Index:	Number of input attributes:	Description:
Old-growth index (Acker et al., 1998)	4	Measures the degree of similarity between old-growth Douglas fir conditions.
Additive index called the LLNS diversity index (Lähde et al., 1999)[<i>The index name is composed of the four authors initials</i>]	8	Distinguishes between the successional stages of Finnish boreal forest.
Biodiversity index (Van Dem Meerschaut and Vanderklove, 1998)	18	Used to characterise biodiversity in Belgian forests. Attributes benchmarked against levels in forest reserves.
Vegetation condition score (Parkes et al., 2003 * ¹ , Oliver and Parkes, 2003 * ² , Gibbons et al., 2004 * ³)	8* ¹ * ² , or 10* ³	Assesses vegetation condition in temperate Australian ecosystems. Attributes benchmarked at the scale of vegetation community.
Rapid ecological assessment index (Koop et al., 1994)	9	Attribute levels benchmarked against levels in unlogged natural forests.
Extended Shannon-Weiner index (Staudhammer and Lemay 2001)	3	Uses an averaging system to extend the Shannon-Weiner index to height, DBH and species abundance.
Index of structural complexity (Holdridge, 1967, cited in Neumann and Starlinger, 2001)	4	Based on traditional stand parameters, which are multiplied together. Sensitive to the number of species.
Stand diversity index (Jaehne and Dohrenbusch, 1997, cited in Neumann and Starlinger, 2001)	4	Combines measures for the variation in species, tree spacing, DBH, and crown size.
Structural complexity index (Zenner, 2000)	2	Measures height variation based on tree height and spatial arrangement of trees.
Structure index based on variance (STVI) (Staudhammer and LeMay, 2001)	2	Based on covariance of height and DBH. Independent of height or DBH classes.
Cantarello and Newton's (2008) index of ecological indicators	17	Used to characterise the ecological condition of UK and Italian woodlands in line with Natura 2000 guidelines. Benchmarked against documented (semi-ancient) broadleaved woodland.
Aggregation index (Clark and Evans, 1954)	2	Describes the horizontal distribution of trees using the relation of the mean distance between the reference tree and its nearest neighbour.
Diameter differentiation index (Füldner, 1995 cited in Vorčáket al., 2006)	3	Tree differentiation within a stand. The index quantifies the differentiation between thinner and thicker DBH between neighbouring trees.

2.2.2.4 Selection of indices to test

These indices vary in terms of number of inputs and computational complexity; for example some indices focus on a single facet of the woodland environment, such as the spacing of tree stems, whereas others utilise multiple facets (up to 17) such as species composition, tree size, size variability, the presence of deadwood and the regeneration of species. It should be noted that some indices require the input of area-based or individual tree-based metrics. Thus, a subset of indices was selected from the horizontal diversity, vertical diversity, and complex diversity index types for use in this project which represent a range of different inputs and computational complexity.

These indices were: (i) the Clark-Evans aggregation (Clark and Evans, 1954); (ii) tree diameter differentiation (Földner, 1995 cited in Vorčák et al., 2006); (iii) vertical evenness (Neumann and Starlinger, 2001); (iv) complexity (Holdridge, 1967); (v) complex stand diversity index (Jaehne and Dohrenbusch, 1997, cited in Newmann and Starlinger); and (vi) the scoring based index developed by Cantarello and Newton (2008). These indices present variation in terms of number of inputs and calculation complexity. Each of these indices will be described in detail in section 4.1.3.

2.2.3 Potential issues with the fieldwork approach

The results of fieldwork are reliant on the judgement of the person carrying out the assessment (JNCC, 2004). Often visual or qualitative assessments are made, rather than quantitative measurements (Newton et al., 2009b). This adds concerns over the subjectivity of data capture, as this will vary for a single person depending on the abilities and bias of the surveyor.

As mentioned previously, it is impossible for the researcher to measure every possible thing within the environment. The scope of measurement of potential indicators and the spatial extents in which to record them must include substantial consideration of costs in terms of both labour and time (Aplin, 2005). Ecosystems are in a constant state of change, fieldwork must be planned in a manner to minimise temporal delay between data collected from different areas in order to be representative of the site conditions as a whole at a given time. Datasets recorded years apart could contain drastically different information.

Not including destructive methods of data capture, fieldwork can potentially be disruptive or damaging to organisms or environments within field plot areas, for example through the potential trampling of vegetation or the unintentional movement of organisms which may colonise a site. Another point to consider is the representativeness of the sample plot to the whole of the area under investigation. Additionally in some situations access to a specific area may be an issue, for example due to inaccessible terrain or land ownership (Campbell and Wynne, 2011).

2.2.4 Assessment of forest condition within this project

Due to the differences in the definitions of woodland condition between different management objectives it is necessary to define what condition is in the context of this research project and to identify the critical environmental metrics necessary in order to assess it.

2.2.4.1 Woodland condition definition

Condition can be defined according to compositional and structural attributes, evidence of natural turnover and anthropogenic influences. For the composition attributes, tree species richness and the presence of native species are considered necessary for condition assessment. For the structural components, the density of individual trees, tree size, and variation in size are considered necessary. Natural turnover refers to the presence of regenerating tree species (seedlings and saplings) and native regenerating trees, in addition to the availability of dead material. Anthropogenic influences such as disturbance and damage are linked to the previously mentioned attributes. However disturbance and damage is difficult to quantify, and therefore condition assessment in this project will focus upon compositional, structural, and natural turnover attributes. Thus, a site which optimises each of these facets (e.g. large tree sizes with high species richness), relative to one another, would constitute a site of high value or good condition.

2.2.4.2 Identification of required field metrics

The previous sections have highlighted the need for and examples of field condition indicators for various forest structural and compositional elements. In reference to objective one of this research project, a listing of forest condition indicators considered important in both the academic and grey literature are presented in Table 2.4. The following sections of this literature review will focus upon identifying examples of airborne remote sensing based

research relevant to the extraction and the estimation of those relevant forest indicator metrics in order to complete objective two.

Condition monitoring requirements in the context of this thesis are therefore defined as methods capable of reporting on each of the various composition and structural attributes reported in Table 2.4. The list of required monitoring attributes satisfies both forestry and ecological information requirements. Therefore, the remainder of this literature review will identify the remote sensing technologies and techniques which offer the potential to satisfy the information requirements of assessing forest condition.

Table 2.4 – A listing of forest condition indicators assessed through field operations

No.	Indicator name	Description	Authors
Stand structure			
1	Stem density	Composition and density will have a primary influence on woodland structure and composition at lower vertical levels via shading and the abundance and the distribution of gaps.	(Cantarello and Newton 2008)
2	Tree size	Tree size is typically quantified in terms of mean DBH and basal area. Biomass can also be calculated. DBH generally increases with stand age and is generally used to discriminate between successional stages in pine and boreal forests. Stand basal area is an attribute directly related to mean DBH. It is also indicative of stand volume and biomass.	(Cantarello and Newton, 2008) (McElhinny et al., 2005) (Vorčák et al., 2006) (Van Den Meersschaut and Vandekerhove, 1998)
3	Number of storeys	Also known as canopy layers or strata. Multi-layered stands are expected to increase the diversity of habitat niches within stands. There are also links to light penetration through the canopy.	(Van Den Meersschaut and Vandekerhove, 1998) (McElhinny et al., 2005) (Neumann and Starlinger, 2001)
4	Canopy closure	Defined as the proportion of the stand covered by live crown, where the crown is often considered as an opaque object. There are also links to light penetration through the canopy.	(Van Den Meersschaut and Vandekerhove, 1998) (McElhinny et al., 2005)
5	Average height (m)	Links have been made between successional stages, the number of vertical strata and stand biomass. Tree height has also been used as a surrogate for tree age/maturity. Older trees are more likely to produce various resources for the local environment.	(Cantarello and Newton, 2008) (McElhinny et al., 2005) (Hinsley et al., 2006)
6	Tree spacing	Stands with lower tree spacing or which have been thinned produce larger trees more quickly than denser, non-thinned stands.	(Horner et al., 2010) (Kim et al., 2009a) (Vorčák et al., 2006)

Table 2.4 (continued)

7	Tree crown dimensions (e.g. crown horizontal area, depth, and height to the first live branch)	Crown dimensions are strongly correlated with stem diameters and, therefore, to forest volume and biomass. Competition between trees within denser stands may reduce the size of the live crown, and by extension, stem diameters. Crown sizes have links to total photosynthetic potential and light penetration through the canopy.	(Kalliovirta and Tokola 2005) (Popescu and Zhao 2008) (Jaehne and Dohrenbusch, 1997, in Newmann and Starlinger, 2001)
Stand composition			
1	Number of tree species	Overstorey trees will influence the availability to the understorey of other resources, such as water and nutrients, and can modify the chemical characteristics of the litter layer, which in turn can affect the diversity of the understorey.	(Van Den Meersschaut and Vandekerhove, 1998) (Hill et al., 2010) (Noss, 1990)
2	Tree species diversity	Species diversity as a composition which refers to the identity and variety of elements in a population includes species lists and measures of species diversity and genetic diversity. An example is the Shannon-Weiner index.	(Cantarello and Newton, 2008) (Van Den Meersschaut and Vandekerhove, 1998)
Deadwood			
1	Volume of downed deadwood	The importance of deadwood for biodiversity in forests is high; this is mainly due to its role in nutrient cycling and species habitation. Thinned timber which has been left on the ground has been found to produce additional biodiversity benefits. Slow decomposing woody material may provide additional carbon storage.	(Cantarello and Newton, 2008) (Pesonen et al., 2008) (Horner et al., 2010) (Van Den Meersschaut and Vandekerhove, 1998)
4	Volume of standing deadwood	Standing dead trees can be an important source for tree hollows which are of importance to species habitation and nutrient cycling.	(Cantarello and Newton, 2008) (Pesonen et al., 2008) (Van Den Meersschaut and Vandekerhove, 1998)
Regeneration of trees			
1	Number and composition of total saplings and seedlings	Composition and density will have a primary influence on woodland structure and composition at lower levels via shading and the abundance and distribution of gaps. Regeneration is mainly of interest for future biodiversity, however because of its sensitivity for silvicultural measures, it is an important parameter for evaluating the impact of forest management.	(Cantarello and Newton, 2008) (Van Den Meersschaut and Vandekerhove, 1998)
Ground Vegetation			
1	Number of vascular plant species, bryophytes species and total ground vegetative cover	The diversity of vascular plants, bryophytes and proportional cover are indicators of biodiversity. Vascular plants host specialised animal species and thus have a link to faunal diversity. Bryophytes can act very quickly to changing environmental conditions.	(Cantarello and Newton, 2008) (Van Den Meersschaut and Vandekerhove, 1998)

2.3 Applications of remote sensing for forest condition assessment

The use of remote sensing data in ecological studies has been relatively limited (Newton et al., 2009b). Owing to the spatially heterogeneous nature of landscapes, ecological approaches can potentially be applied at a variety of scales to a wide range of different environments. Ecology involves the investigation of organisms in their environmental setting. These investigations require spatially explicit data. Traditionally ecological data are collected through manual field observations, as described in sections 2.1 and 2.2. The benefit of this approach can be high accuracy because of its labour-intensive nature, but it is generally impractical for anything other than local scale studies (Aplin, 2005; Vierling et al., 2008).

However, ecological analysis can extend well beyond the local scale and there is considerable need for, and interest in, ecological investigation at wider spatial scales (Newton et al., 2009b; Turner et al., 2003). Remote sensing techniques have consequently become more frequent in many ecological investigations, providing the only realistic, cost-effective means of acquiring data over large areas (Ker and Ostrovsky, 2003). Airborne and spaceborne remote sensing technologies are recognised as powerful tools to acquire detailed and synoptic data on various landcover types. To date however, remote sensing has seen limited uptake in the discipline of ecology. The role of remote sensing in landscape ecology might be strengthened by greater integration of diverse remote sensing data with a broader range of ecological data (e.g. beyond species presence/absence), and by increased recognition of the value of remote sensing beyond landcover mapping or pattern description (Newton et al., 2009b).

The reporting of forest habitat status requires detailed knowledge of many aspects of habitats at multiple spatial levels. The techniques commonly used for assessing forest environments in the UK for multiple land uses are field survey and aerial photography (Falkowski et al., 2009b). Aerial photography is chiefly used by the Forestry Commission for inventory purposes, commonly through manual interpretation of the imagery. Both of these techniques have distinct disadvantages for determining measures of vegetation structure, composition, and spatial metrics at larger (landscape) scales.

Ground surveys utilise sampling designs to select plots where details of woodland condition are recorded; these parameters are then extrapolated up to stand-level and above (Wang et al., 2004). Ground survey information is typically detailed, covering many attributes such as vegetation types, structure, abundances, dead materials and assessment of vegetative growth (Cantarello and Newton, 2008). Regardless of sampling design, ground survey is expensive, time consuming and limited by those features encountered in the sample plots (MacArthur and MacArthur 1961; McRoberts et al., 2003). Additionally, many vegetation structural elements are also often out of reach of ground sampling.

Aerial photography however can cover vast areas, in the order of kilometres, in a fraction of the time fieldwork can be accomplished. Measurements from air photos are made manually, which is time consuming and prone to being inconsistent between interpreters (Culvenor, 2002; Koukoulas and Blackburn, 2005a). Another major limitation is the inability to see through the tree canopy, thus only inferences of tree height and structure can be made.

Monitoring, conservation and management of forests are hindered by the lack of spatially and temporally extensive information on tree species and floristic composition, and are further hindered by prohibitive costs, site inaccessibility and lack of the ability to measure attributes in an objective and repeatable manner. Additionally measurements such as vertical structure, biomass and species composition are difficult to extrapolate to the landscape scale because of the complexity of the system (Lee et al., 2004). Remote sensing technologies have facilitated considerable advances in the modelling, mapping, and understanding of ecosystems. Typical applications involve either aerial photography or satellite imagery, such as that of Landsat Thematic Mapper or to a lesser degree active radar sensors (Lefsky et al., 2002). These types of sensors have been proven to be satisfactory for many ecological applications, and enable researchers to analyse spatial patterns. Although Synthetic Aperture Radar (SAR) offers promise for identifying sparse vegetation and mapping vegetation in floristically simple landscapes, this technique is too insensitive to high biomass levels, and is therefore unsuitable for mapping some types of forest vegetation (Rignot et al., 1994; Smith et al., 1994; Waring et al., 1995).

However, conventional sensors have significant limitations for ecological applications, in terms of the spatial resolution and the amount of information which can be extracted. The sensitivity and accuracy of both spectral and Radar devices have been shown to fail with increasing above-ground biomass and leaf area index (LAI) levels. Optical data in particular is limited to only representing two-dimensional spatial patterns.

With the advent of fine-scale spectral and spatial resolution remote sensing technologies alternatives for the retrieval of forest floristic and structural information can be realised (Lucas et al., 2008a). The presence of specific organisms and overall species richness is strongly dependent on the three-dimensional spatial pattern of vegetation. For example, bird species in particular are often associated with specific three-dimensional structural features within forest environments (Hinsley et al., 2002; 2009a). In addition, such remote sensing capabilities can reduce the amount of field work necessary and fulfil the needs of special purpose surveys (Leckie et al., 2003b).

To date many of the remote sensing studies within the academic literature have focused upon specific issues or the extraction of a selection of attributes, where not all of them are relevant to the assessment of forest condition. Only a limited number of studies have been aimed at assessing forest condition. A number of recent studies have attempted to map one of the key components of condition, for example the volume of standing and fallen deadwood (Mücke et al., 2012), or have attempted to assign a basic three-tier classification through the identification of species richness, vegetation height and patch connectivity (Simonson et al., 2013). Thus they have been limited to a single potential component or to a relatively coarse scale with a basic classification

What follows is an overview of the surrounding remote sensing literature, with examples of research into extracting metrics identified as indicators of forest condition utilising two of the most promising digital data sources, namely passive optical spectral sensors and active LiDAR sensors.

2.3.1 Multispectral and hyperspectral sensors

Hyperspectral sensors are a passive sensor technology whereby images of a scene are collected in tens to hundreds of narrow spectral bands nearly simultaneously (Campbell and Wynne, 2011). They represent the next step in the spectral dimension of the evolution of multispectral imaging radiometers. The term hyperspectral usually refers to an instrument, whose spectral bands are constrained to the region of solar illumination, i.e., visible through shortwave infrared and in the remote sensing context has an observing platform that is either airborne or spaceborne. The data collected are often termed an ‘image cube’ where the two spatial dimensions are joined by a third ‘spectral dimension’.

Multiple airborne sensor designs exist, such as optical-mechanical scanners and charged-coupled devices (Campbell and Wynne, 2011). Two examples of well-established aircraft mounted sensors are the Compact Airborne Spectrographic Imager (CASI) and Airborne Visible/Infrared Imaging Spectrometer (AVRIS) (Lucas et al., 2008a). A review of the specific technology is beyond the scope of this project, but a comprehensive guide is available in Campbell and Wynne (2011). These sensors are coupled with differential Global Positioning Systems (GPS) to provide an aircraft inertial measurement data; this is required in later stages when the imagery is processed (Campbell and Wynne, 2011).

There are two main optical sensor types: multispectral and hyperspectral. The multispectral sensor records reflectance at the sensor through several broadly defined spectral channels, typically 3 to 15 (Campbell and Wynne, 2011). Hyperspectral sensors expand on the capacity of multispectral sensors, by allowing the detection of a greater number of spectral bands, which can have hundreds of narrowly defined spectral channels (Lefsky et al., 2001). Spatial resolution for both systems can vary between centimetres and tens of metres depending upon the device and operating platform.

High spatial resolution imagery in which individual trees can be separated is necessary for forestry management; current airborne optical sensors can potentially perform this role in terms of spectral and spatial resolution (Clark et al., 2005; Leckie et al., 2003a). Success in a number of fields has been presented using hyperspectral sensors for determining species-level abundance patterns in a variety of biomes (Zhang et al., 2006).

The following sections identify those multi or hyperspectral remote sensing case-studies which extract or estimate the forest condition indicator metrics as stated in Table 2.4.

2.3.1.1 Forest structure

The majority of studies which estimate biophysical variables from multi and hyperspectral remotely sensed data have used empirical techniques to relate spectral data and various derivatives to biophysical parameters. If biophysical parameters are strongly correlated with remotely sensed radiance data, then these data can be used to predict those biophysical characteristics over large areas (Treitz and Howarth, 1999).

Franklin et al. (2001) utilised a hierarchical classification to identify species composition in some 30 stands in New Brunswick, Canada. Utilising multispectral data from a CASI-2 sensor, stands of differing species composition, crown closure and stem density were classified using a combination of parcel-based texture analysis and standard pixel-based nearest-neighbour spectral classification algorithms. Classification accuracy increased from 54% to 75% by the inclusion of image texture. Influences of shadowing in stands of a different structure create different texture patterns in different stem densities, even if they have the same species composition. Airborne spectral data can be used to derive surrogates for stand structure, but cannot supply direct measurements of canopy height and surface roughness, or estimates of tree stem diameter or timber volume.

Optical sensors with high enough spatial resolution have been demonstrated to estimate canopy characteristics of crown diameter, crown closure, and stand density. In some cases automated methods have been utilised to extract individual tree crown (ITC) information from optical imagery, such as that used in Bunting and Lucas (2006) in Queensland, Australia. ITCs were delineated using a spectral index created from CASI multispectral bands to discriminate forest and non-forest vegetation, known as the forest discrimination index (FDI) and an automated watershed segmentation technique was used to identify tree crowns and tree clusters. Tree locations, species and crown dimensions were then extracted. This approach provided accuracies of $> \sim 70\%$ for individual trees and clusters of trees of the same species, with lower accuracies associated with denser stands containing multiple canopy layers. Similarly automated methods can be applied to delineate forest stands (Leckie et al., 2003b).

The vast majority of studies which estimate biophysical variables from optical data have used empirical techniques to relate spectral data to biophysical parameters. Remotely sensed radiance data can be used to predict those biophysical characteristics over large areas. Spectral index-based approaches have been used to estimate vegetation parameters such as LAI and biomass (Asner, 1998).

Plant health and vigour have been examined through the use of high spectral resolution remote sensing, whereby sensitive features within the returned spectra are identified, such as the ‘green peak’ and ‘red-edge’ which are located within the spectral curve (Treitz and Howarth, 1999). Examples of insensitive features can correspond to the ‘chlorophyll well’ or the ‘near-infrared plateau’ within the spectral curve for vegetation responses. Broad waveband sampling, such as that in many multispectral systems, can easily combine the reflectance from narrow sensitive and insensitive features, masking out the response of sensitive features. Developments in hyperspectral sensors have reduced this problem (Lefsky et al., 2001).

Reflectance in the visible (400-700nm) and the near-infrared (700-2500nm) portions of the electro-magnetic spectrum is most important in the detection of vegetation stress, since changes in leaf chlorophyll and/or water content significantly affect leaf reflectance (Treitz and Howarth, 1999). Work by Pu et al. (2008) utilised CASI hyperspectral imagery to detect mortality and vegetation moisture stress caused by hardwood oak disease, suspecting that the foliage of the infected trees, even if appearing green, had a different water and biochemical status as compared with healthy leaves. The spectral difference between healthy and stressed oak leaves may be slight however. Given a traditional classification strategy, the separation of healthy from non-healthy trees was almost impossible. Using a multilevel classification and principle component analysis for the visible region and NIR bands, living and dead vegetation could be determined to 80% accuracy, while separation of non-stressed oak trees was accurate to 76%. Levels of accuracy were influenced by levels of shade and shadow, in addition to the timing of imagery acquisition.

Airborne optical remote sensing has well-developed and understood techniques for vegetation mapping, following the development of technologies and techniques over the past 30 years for forest ecosystem management applications. This development has produced mature techniques ready for implementation in the management of forest resources. Examples of established techniques include spectral data transformations, typically vegetation indices, often used to identify vegetation (Bork and Su, 2007).

A common assumption with spectral indices is that the transformations of spectral band reflectance are more closely correlated with plant biophysical qualities and are generally less sensitive to external variables such as solar zenith angle (Treitz and Howarth, 1999). A frequently used index is the normalised difference vegetation index (NDVI) contrasting optimum reflection and absorption characteristics, although other indices assist in vegetation and soil identification (Bork and Su, 2007), often being applied to satellite or large scale multispectral imagery. Along with the NDVI, the most common vegetation indices utilise the information content of the red and near-infrared canopy reflectance, being highly correlated with green-leaf biomass. Table 2.5 lists the most commonly referred to spectral indices cited in the surrounding literature.

Estimating leaf area index (LAI) from remote sensing relies upon the unique spectral response of green leaves in contrast to other land surface materials. LAI is a standard expression for leaf area of a plant community and is generally defined as the total leaf area per unit of ground cover. LAI has been found to have close associations with light interception, gas exchange, carbon flux, photosynthesis and biomass production (Treitz and Howarth 1999; Zheng and Moskal 2009). Zheng and Moskal (2009) mention that a number of spectral indices have been used to extract estimates of LAI, typically from passive satellite remote sensing. In addition to NDVI, the enhanced vegetation index (EVI), and reduced simple ratio (RSR) have been used. One major issue of retrieving LAI from vegetation indices based on the different band combinations from multi or hyperspectral remote sensing is saturation at high LAI, which means the relationship between an index and LAI will not increase linearly (Zheng and Moskal, 2009).

Another use of spectral indices is the estimation of chlorophyll concentrations within the forest canopy, again utilising bands in the narrow visible and near infrared (Dash and Curran, 2004). Chlorophyll is an important foliar chemical and the content within a canopy is related positively to both the productivity of that vegetation and the depth and width of the chlorophyll absorption feature in the reflectance spectra. The position of this spectral absorption feature influences the position of the red edge feature, moving it to longer wavelengths. Thomas et al. (2008) investigate twenty such indices which have demonstrated strong relationships to chlorophyll content in complex mixed wood forest canopies. Each of these indices was combined with a LiDAR derived canopy height model and compared using regression analysis. Many of those indices proved to be insufficient in mixed forest, but proved effective in single species forest. With the inclusion of height data the relationship between index and chlorophyll was strengthened.

Spectral data have been used to model the movement of different animal species and foraging dynamics, for example Mongolian gazelles (*Procapra gutturosa*), using NDVI-based estimators, derived from 16 day composite satellite MODIS imagery. Such measures of landscape resistance based on habitat characteristics could help to explain gene flow in the species (Mueller et al., 2008). By estimating local environmental conditions, researchers have been able to predict the likelihood of animal species presence (Geffen et al., 2004; Pilot et al., 2006). Thus, spectral data could be a convenient way of characterising individual plot-scale areas or entire landscapes in terms of their environmental characteristics, and ultimately to aid studies in adaptive landscape genetics.

Table 2.5 – A list of the most common vegetation indices derived from spectral imagery

Index:	Description:	Authors:
Greenness – High pixel values indicate high proportions of green biomass		
Normalized Difference Vegetation Index (NDVI)	Normalized difference of green leaf scattering in near-infrared, chlorophyll absorption in red wavelengths.	(Jackson et al., 1983; Rouse et al., 1973; Sellers, 1985; Tucker, 1979)
Simple Ratio Index (SRI)	Ratio of green leaf scattering in near-infrared, chlorophyll absorption in red wavelengths.	(Rouse et al., 1973; Tucker, 1979; Sellers, 1985)
Red Edge Normalized Difference Vegetation Index (RENDVI)	A modification of the NDVI using reflectance measurements along the red edge.	(Gitelson and Merzlyak, 1994; Sims and Gamon, 2002)
Vogelmann Red Edge Index 1 (VOG1)	A shoulder of the RED-to-NIR transition that is indicative of canopy stress.	(Vogelmann et al., 1993)
Red Edge Position Index (REP)	The location of the maximum derivative in near-infrared transition, which is sensitive to chlorophyll concentration.	(Curran et al., 1995)
Light Use Efficiency – A measure of the efficiency with which vegetation is able to use incident light for photosynthesis.		
Photochemical Reflectance Index (PRI)	Useful to estimate absorption by leaf carotenoids (especially xanthophyll) pigments, leaf stress, and carbon dioxide uptake.	(Gamon, 1992)
Structure Insensitive Pigment Index (SIPI)	Indicator of leaf pigment concentrations normalized for variations in overall canopy structure and foliage content.	(Penuelas, 1995)
Dry or Senescent Carbon – Provides an estimate of the amount of carbon in dry states of lignin and cellulose		
Normalized Difference Lignin Index (NDLI)	Detects leaf lignin increases at the 1754 nm feature relative to the 1680 nm canopy structure region.	(Serrano et al., 2002)
Cellulose Absorption Index (CAI)	Detects absorption features due to cellulose above 2000 nm wavelength.	(Daughtry et al., 2004)
Leaf Pigments – Designed to provide a measure of stress-related pigments present in vegetation.		
Carotenoid Reflectance Index 1 (CRI1)	Detects a relative difference in absorption indicative of changes in leaf total carotenoid concentration relative to chlorophyll concentration.	(Gitelson et al., 2002)
Anthocyanin Reflectance Index 1 (ARI1)	Changes in green-wavelength absorption relative to red-wavelength absorption indicating leaf anthocyanins.	(Gitelson et al., 2001)
Canopy Water Content – Applications include canopy stress analysis, productivity prediction and modelling, fire hazard condition analysis, cropland management, and studies of ecosystem physiology		
Water Band Index (WBI)	Absorption intensity at 900 nm increases with canopy water content.	(Penuelas et al., 1995)

2.3.1.2 Forest composition

The developments in high spatial resolution airborne hyperspectral imagery have allowed greater discrimination of tree species when compared with data collected by previous coarser (>5m) spatial resolution multispectral sensors (Lucas et al., 2008a). The discrimination of species through image spectra however remains complicated as canopy elements can obscure each other in the overstorey, the reflectance of many vegetation species is similar, and there is often as much spectral variability within tree crowns of the same species as between species (Lucas et al., 2008a).

Many studies have utilised automated techniques to classify different surface materials, such as vegetation, water and man-made structures. The most dominant of these methods is that of per-pixel classification, where a defined set of spectral characteristics is set as a particular landcover class. More recently object-oriented approaches have been explored, for example detection of a group of pixels conforming to a tree crown, and examining the combined values of the pixels within the group.

A number of studies have succeeded in the discrimination of and/or mapping of individual tree species. Many of these have focused upon temperate forested areas within northern latitudes, where forests are generally of simpler structure and contain fewer species types. Leckie et al. (2003b) for example, attempted to discriminate between seven tree vegetation classes in British Columbia, Canada, using data acquired from a CASI sensor. The imagery used had spatial resolution sufficient to delineate individual tree crown information. Thus applying an appropriate automated individual tree crown (ITC) delineation algorithm each tree crown could be identified and assigned a species classification. This research also proposed the measurement of spectra from the best illuminated region within each crown in an effort to account for differing illumination. By aggregating the ITCs, counts of each species within each stand were produced.

Similarly Lucas et al. (2008a) used an ITC delineation algorithm, and assigned classifications based upon the reflectance spectra within those areas identified as the tree crowns, to produce species maps at the individual tree crown/cluster level. The study was performed in central Queensland, Australia, using CASI and HyMap sensors acquired in the autumn of 2000. The standard visible to near infrared spectra recorded from the CASI data were found to be insufficient in many cases for accurate species classification, but the inclusion of shortwave

infrared spectra recorded by the HyMap system improved classification accuracies for most species.

Mapping tree species in optical imagery is a difficult task as spectral variation within species is typically higher than between them. As a result considerable variation has been reported in the accuracy of species mapping in single date optical imagery (Hill et al., 2010). This is especially true in the case of deciduous species. The problem can be at least partly overcome by the use of time-series data, which can capture phenological change such as leaf flush or senescence and are likely to increase the spectral separability of deciduous tree species, ideally across a single growing season. Hill et al. (2010) investigated the ability to create an overstorey tree species map using five dates of airborne multispectral imagery, which captured information over a single growing season. The study compared the classification of six tree species on individual images and different combinations of images. One of the single images achieved a classification accuracy of 71%. However higher accuracies were achieved when combining three of the images producing an overall accuracy of 88% after processing.

For individual tree-level characteristic estimation for inventory, species typing is required and this is potentially best provided by hyperspectral data (Lucas et al., 2008b). While readily providing information on the horizontal organisation of vegetation canopies, vertical structure has generally proven difficult to obtain (Bork and Su 2007; Lucas et al., 2008a). Knowledge of the vertical distribution of canopy elements may enhance species classification accuracy, in addition to aiding in understanding a number of ecosystem processes. Vertical structure has been indirectly assessed through the influence of variation in shadowing with changing density, as with photogrammetry, or by the identification of spectrally unique scene components within the canopy, such as lichen (Bork and Su, 2007). The reflectance of the vegetated canopy is influenced by shadowing within and between tree crowns. This varies with tree shape, structure, density and the relative positions of individuals in the vertical profile (Bunting and Lucas, 2006), complicating the delineation procedure.

Finding ecological proxies of species diversity is important for developing effective management strategies and conservation plans over multiple scales. As mentioned previously, biodiversity assessment has traditionally been carried out via fieldwork. The causal relationship between species diversity and environmental heterogeneity has been of long-standing interest amongst ecologists, when considering that areas with higher environmental heterogeneity can host more species due to a greater number of available niches (Rocchini et al., 2010a). The use of remote sensing for estimating environmental heterogeneity and hence species diversity is a powerful tool as it can provide a synoptic view and cover large areas.

Spectral variability, or heterogeneity, in the remotely sensed signal is expected to be related to environmental heterogeneity and could be used as a proxy for species diversity (Oldeland et al., 2010; Rocchini et al., 2010a; Rocchini et al., 2010b), with correlations with landscape structure and complexity. Palmer et al. (2002) state the spectral variation hypothesis as that of the greater the habitat heterogeneity, the greater the species diversity within it, regardless of taxonomic group under consideration. Rocchini et al. (2010a) report that the higher the spectral resolution (number of bands), the higher the power to discriminate objects which reflects in a different manner, or the power to detect an area's heterogeneity. Different vegetative species for example will respond differently to light in the electromagnetic spectrum. It is more efficient to identify those spectral bands which really matter in terms of diversity for research efforts. Further details of this can be found in Rocchini et al. (2010a). It is important to note that the measure of spectral heterogeneity being applied depends on the final goal of a study.

2.3.1.3 Attribute summary

From the above discussion it is clear that spectral remote sensing, primarily hyperspectral, can provide a wealth of information for the modelling, definition and mapping of spatial, biophysical and temporal patterns in forest ecosystems. In terms of attributes relevant to the assessment of forest condition, research into hyperspectral derived metrics indicates that many overstorey structural and composition components may be extracted or derived. However metrics relating to tree size and heights of canopy components cannot be achieved. The estimation of elements of the understorey, such as deadwood, regenerating species and ground vegetation metrics are also underrepresented here.

2.3.2 Light Detection and Ranging (LiDAR)

Light Detection and Ranging (LiDAR) it is an active remote sensing technology that determines distance or range, by taking the product of the speed of light and the time required for an emitted laser pulse to travel to and return from a target object (Lim et al., 2003b). The elapsed time from when the laser pulse is emitted to when it is received can be measured and used to identify the three-dimensional position of each range point.

LiDAR is quickly gaining prominence in natural resource research and management due to its inherent ability to represent complex vertical structures and ground surfaces with very high precision (Evans et al., 2009). LiDAR is capable of providing both geo-referenced horizontal and vertical information on the structure of forest canopies with sampling dependent on the type of LiDAR system used (Evans et al., 2009).

Multiple LiDAR system designs exist which have applications within forested environments. What follows is a brief description of the two main groups, scanning and profiling.

2.3.2.1 Types of airborne LiDAR systems

The LiDAR active sensor design uses pulsing lasers which generate very specifically timed bursts of light. There are two main forms of airborne LiDAR system: profiling, discrete-return and waveform. The first of these, profiling lasers, refers to the application of lasers aimed directly beneath the aircraft to illuminate a single region in the nadir position. When these devices are used to acquire topographic data they are known as ‘airborne laser altimeters’. The forward motion of the aircraft carries the illuminated region forward to view a single track directly beneath the aircraft. The echoes from the repetitive LiDAR pulses provide an elevation profile of the narrow region beneath the aircraft platform (Campbell and Wynne, 2011). This system does not generate imagery in the traditional sense, but rather spot measurements along the flightline.

Profiling laser systems usually operate with a large footprint; this refers to the area illuminated by the laser pulse. A large footprint typically refers to illumination of an area with a diameter $> 1\text{m}$. Such systems can use continuous wave lasers, which generate a continuously modulated beam of light. The instrument transmits this radiation to the ground and receives a digitised representation waveform for each pulse, providing vertical profile measurements (Coops et al., 2004). Each waveform consists of a series of temporal modes

(called echoes), where each of them correspond to an individual reflection from an object, and sometimes with the ability to discern changes in signal amplitude, phase and intensity.

A scanning LiDAR system utilises a scanning mirror which directs the laser pulses back and forth across a horizontal swath beneath the aircraft. Such a scanning system can transmit upwards of 2,000 pulses each second, thus as the aircraft moves forward points are captured either side of the aircraft's position, building up measurements from multiple swaths (Campbell and Wynne, 2011). Scanning LiDAR typically use wavelengths in the near-infrared region of the spectrum ($\sim 1.64\mu\text{m}$) due to its sensitivity to vegetation features and freedom from atmospheric scattering. LiDAR data are typically acquired in parallel strips that stitch to form a continuous area of coverage. Scanning LiDAR systems record discrete-return data, and due to recent technological improvements waveform data can now also be recorded. Scanning LiDAR can operate using discrete-return or waveform techniques.

Discrete-return LiDAR (DR) system operate using a small footprint which may acquire data from areas as small as 0.15-0.60m in diameter (Campbell and Wynne, 2011). The criterion for collecting multiple returns is based on the intensity of the laser energy returned to the sensor, with up to five returns per laser pulse systems being most common. The DR system records one to several returns through the vegetated canopy, in a vertically non-systematic manner (Evans et al., 2009).

More recent developments in sensor design have allowed the advance of small-footprint and continuous wave LiDAR, in order to produce a 'hybrid' of the two technologies, this is known as full-waveform (FW) LiDAR. FW sensors digitize the total amount of energy returned to the sensor in fixed distance intervals, providing a continuous distribution of laser energy for each laser pulse (Evans et al., 2009). Instead of clouds of individual three-dimensional points, such as with DR LiDAR, FW devices provide connected profiles of the 3D scene, which contain more detailed information about the structure of the illuminated surfaces (Alexander et al., 2010; Mallet et al., 2009; Miura and Jones 2010). Each waveform consists of a series of temporal modes (called echoes), where each of them correspond to an individual reflection from an object, in much the same way as large-footprint continuous wave systems. Differences in signal amplitude, phase and intensity can also be derived.

DR systems suffer from a sizable ‘blind spot’ following each detected return, during which no other returns can be detected (Reitberger et al., 2008). This blind spot can be 1.2-5m, because of limitations in the sensor design. FW systems remedy this problem through post-processing the waveform data to identify proximal peaks which would otherwise be treated as one, for example Chauve et al. (2007) resolved an additional 40-60% of returns from FW data in comparison to DR. The most common approach is to approximate the waveform as a series of Gaussian curves; fitted by a non-linear least squares approach (Wagner, 2006). FW LiDAR datasets can also provide additional information about the structure of the illuminated surfaces from interpretation of the backscattered waveform (Chauve et al., 2007). Amplitude, echo-width and the greater number of returns are examples of additional information obtained from waveform data.

2.3.2.2 Applications of LiDAR remote sensing for ecology

Lefsky et al. (2002) state that only a few areas of application for LiDAR remote sensing have been rigorously evaluated, while many other applications are considered feasible but have not yet been explored. The developments within the field of LiDAR are occurring so rapidly that it is difficult to predict which applications will come to the fore in the years to come. Scanning LiDAR systems provide spatially intensive and extensive canopy height measures that could facilitate forest inventory at much finer spatial scales than the basic stand operating unit. To estimate forest structural attributes of interest besides canopy height, the LiDAR height measures must be related to field measures of these attributes, recorded in field plots distributed across the full range of variation (Hudak et al., 2008a). Currently, applications of LiDAR remote sensing for forest ecology fall into three broad categories: (i) remote sensing of ground topography; (ii) measurement of the three-dimensional structure and function of vegetation canopies; and; (iii) the prediction of forest structure attributes (such as above-ground biomass).

The following sections identify those DR and FW remote sensing case-studies which extract or estimate those forest condition indicator metrics as stated in Table 2.4.

2.3.2.3 Forest structure

The mapping of topographic features is the largest application area of LiDAR remote sensing, due to its use in land surveys and producing accurate measures of elevation to fulfil a variety of objectives. Within the realm of ecology, the topography of an area both inside and outside of the forest often has a strong influence on the structure, composition and function of ecological systems (Lefsky et al., 2002). For example, many species are constrained by microhabitats resulting from changes in altitude. In addition, the spatial distribution of elevation is a determinant of watershed flow (Turner et al., 2003). Traditional survey and photogrammetric methods are limited because of the lack of characterisation of 3D surfaces (Hudak et al., 2009; Lefsky et al., 2002). LiDAR can penetrate the tree canopy layer providing height information from beneath and within the canopy.

Vegetation height metrics

The most obvious vegetation measure extracted from LiDAR is that of canopy height information which is commonly used to describe forest structure. Calculation of canopy heights requires the identification of a ground reference level, allowing canopy heights to be calculated by subtracting the difference between those LiDAR returns for the canopy and a surface representing the terrain (Lim et al., 2003b). Once the LiDAR point cloud is filtered and the ground returns identified, a digital elevation model (DEM) can be produced for the ground surface by interpolating between these classified points, producing the digital terrain model (DTM). DEMs typically take the form of raster data. The highest elevations from the LiDAR points within a defined grid extent, i.e. raster resolution, can be used to produce a digital surface model (DSM) of the tree crown. The DTM elevations can be subtracted from the DSM to remove the effects of the underlying terrain and produce a normalised digital surface model (nDSM). This is also known as a canopy height model (CHM). Such an approach was employed in Patenaude et al. (2004) for Monks Wood, UK, where the nDSM allow accurate estimate of canopy height to be made.

Vegetation height measurements can be made accurately even on vegetation of short stature (~1m), at least in areas of relatively flat terrain (Lefsky et al., 2002). Lucas et al. (2008a) state that LiDAR data have been shown to produce estimates of tree height that are considered to be at least equivalent to and often more accurate than those obtained at ground-level using traditional approaches, for example with the use of a clinometer. This occurs partly because of the difficulty with ground based measurements in sighting the top of tree crowns that are

expansive or occur where the canopy is dense. However LiDAR derived tree top heights can be less accurate, albeit often to a lower level than ground survey, for trees where the crown tapers to a point, such as in many conifer species, unless a very high sampling density LiDAR system is used (Lovell et al., 2005). Gaveau and Hill (2003) put forward that LiDAR pulses penetrate into the shrub and tree canopy before being returned resulting in an underestimation of height. The penetration of the upper canopy by each laser pulse will vary with small scale variation in closure of the upper canopy surface (i.e. the gaps at a spatial resolution corresponding with laser footprint size) as with variation with leaf area, density, reflectivity, and orientation. This is especially the case in broadleaved deciduous tree species.

As with optical data, LiDAR can be used to estimate planimetric cover and additionally to facilitate retrieval of cover at different layers within the forest itself (Lucas et al., 2008b). LiDAR data can provide densely spaced canopy height measures that can be empirically related to field measures of stand height and other measures of stand structure. The focus upon tree canopy height has been due to its link to predicting variables for other forest attributes, such as biomass and volume. Other studies have used canopy height and the presence of shrub layers as a proxy to estimate tree maturity (Broughton et al., 2006; Hinsley et al., 2006). Information regarding light gaps can be explored using canopy height distributions (Koukoulas and Blackburn, 2004). Many studies for forest based applications for LiDAR have focused upon calibrating and validating the accuracy of LiDAR attribute estimates against those recorded in the field through statistical analysis, where features and predictors are used directly for forest parameter estimation (Hyypä et al., 2004).

Stand structure metrics

Lindberg et al. (2012) outline a method to analyse both DR and FW LiDAR data for the estimation of vegetation volume (m^3) for coniferous and deciduous forest in south-west Sweden. The vegetation volume profile was thus described as the volume of all tree crowns and shrubs in 1 dm [10cm] height intervals in a field plot and total vegetation volume as the sum of the total vegetation volume profile in the field plot. LiDAR profiles were developed from the distribution of 3D points above the ground in 1 dm intervals. Estimates of vegetation volume were then performed using the ratio of vegetation classified points against the total number of points with a log-linear model based on Beer-Lambert law and then rescaled. The Beer-Lambert law relates the absorption of light to the properties of the material through which the light is travelling. The best result was the normalised FW waveform (RMSE

27.6%) whilst the worst was the DR points (RMSE 36.5%). The results suggested that FW data could predict vegetation volume somewhat more accurately, especially so when corrections were applied for the shielding effects of higher vegetation layers.

Naesset (2002) presents a method of estimating forest stand characteristics from airborne LiDAR data for three types of forest using 144 sample field sites – each 200m² in size, which were classified as young, mature (poor condition) and mature (good condition), within a forest site in Våler, Norway. The stand characteristics of mean tree height, dominant height, mean DBH, stem number, basal area and timber volume were estimated through regression analysis using various statistical attributes extracted from the LiDAR data. The assessed accuracies of the predicted metrics were calculated using RMSE as: mean tree height (0.06m), dominant height (0.07m), mean DBH (0.12cm), stem number (0.31ha⁻¹), basal area (0.19m²ha⁻¹) and timber volume (0.19m³ha⁻¹).

A number of studies have explored the extraction of forest attributes at the stand and plot scale using airborne LiDAR. More recently however research has been performed concerning the delineation of individual trees. These categories relate to forestry information requirements whereas from a methodological point of view it is better to divide them into statistical and image processing-based retrieval methods. In the statistical methods, features and predictors are assessed from the laser derived digital elevation models (DEM) and point clouds which are used directly for forest parameter estimation, typically using regression or discriminant analysis (Hyypä et al., 2004).

The use of height percentiles has become a method of estimating forest structural components. The height percentile refers to the distribution of canopy heights and has been used as predictors in regression analysis or non-parametric models for the estimation of canopy density, mean tree height, basal area and volume for the relevant plot or stand (Lefsky et al., 2002; Lim et al., 2003a; Lim et al., 2003b; Maltamo et al., 2005; Naesset 1997a, 1997b, 2002; Naesset and Okland, 2002). Basal area, or cross section of the tree trunk, is calculated from measured tree diameters, which is then summed and divided by plot area (Hudak et al., 2008a). Measurements are usually made for 1 hectare of land for comparison purposes to examine a forest's productivity and growth rate. Forest attributes are then estimated using canopy height and canopy density metrics and generally conform to the proportion of LiDAR returns recorded at the 0, 10, 20...to...90 percentiles of the height distributions to the total

number of pulses. These values are often recorded alongside distributional measures, such as the maximum and mean values, variance, skewness, kurtosis and coefficients of variation (Hudak et al., 2009).

Stand structure attributes and above-ground biomass have been successfully estimated across different forest biomes using a variety of statistical techniques. For example at the stand scale Naesset (1997b) used airborne LiDAR data to produce estimates of timber volume for two Norwegian forests with relatively sparse point densities of approximately one laser pulse every three metres-squared. A multi-regression analysis was used to develop models for stand volume related to predictor variables derived from the LiDAR data. The initial three predictor variables were: mean stand height, mean height of all laser pulses within a stand and mean canopy cover density. The latter is simply the number of tree canopy hits divided by the total number of transmitted pulses. A relationship was found between total vegetation volume, LiDAR stand mean height and mean laser canopy cover density. This resulted in correlation coefficients (R^2) in the range between 0.472 and 0.838 for the two sites. Naesset (1997b) concludes by stating that variables such as stand age and tree species will influence the model produced.

Anderson et al. (2006) utilised LiDAR-derived estimates of several critical forest inventory parameters, in a study carried out in Capital Forest, Washington state, USA. Using 99 field plots, many strong regression relationships were observed between basal area ($R^2 = 0.91$), stem volume ($R^2 = 0.92$), dominant height ($R^2 = 0.96$), biomass ($R^2 = 0.91$) and LiDAR metrics. The same approach was used to estimate several important canopy fuel indicators using LiDAR data (Andersen et al., 2006). Estimates of canopy height, canopy bulk density, and total canopy fuel weight at the stand-level were used for input into wildfire simulation models. Previous estimates were based on plot-scale fieldwork data collection; unfortunately this was unable to capture the variability in stand structure at larger scales. Predictor variables of maximum and mean height, coefficient of variation, several percentile-based metrics and a canopy density metric were regressed against field inventory measures to estimate canopy fuel input variables, resulting in strong relationships being observed.

The random forest (RF) classification algorithm is a statistical classification and regression tree technique, which can also be utilised for variable selection, interaction detection, clustering, etc. The RF algorithm has achieved excellent results in classifying LiDAR remote sensing datasets (Falkowski et al., 2009a; Hudak et al., 2008a; Hudak et al., 2008b). The iterative nature of RF affords it a distinct advantage over other statistical methods. RF grows many classification trees based on the input data, and requires a training dataset. Each forest or tree object will be assigned a class based upon the number of classification tree ‘votes’ accrued. RF has been utilised to provide pixel and plot-scale estimates of forest structure such as basal area, tree density and even tree species (Hudak et al., 2008a; Hudak et al., 2008b), and classification of up to six forest succession stages (Falkowski et al., 2009a) using derived LiDAR metrics. As many as 60 input variables can be entered into the RF algorithm, examples include: topographic slope; heights of various percentiles; intensity range; mean and standard deviation; percentage of vegetation returns at discrete height ranges; etc.

Forest canopy metrics

Leaf Area Index (LAI) is a key forest structural characteristic that serves as a primary control for exchanges of mass and energy within a vegetated ecosystem (Jensen et al., 2008). Most previous attempts to measure LAI from remotely sensed data have relied on empirical relationships between field measures and various spectral indices derived from optical imagery. However as biomass within an ecosystem increases, accurate LAI estimates are more difficult to quantify (Zheng and Moskal, 2009). Jensen et al. (2008) utilised a combination of airborne LiDAR and SPOT-5 imagery derived indices to estimate LAI values over conifer forest stands in the Northern Rocky Mountains, USA. The results from the SPOT-5 data were poor for the study area, explaining less than 50% of the LAI observed in the field, while the LiDAR LAI estimated from the plot based metrics, such as height percentiles, accounted for a higher amount of variation within both of the study sites, with R^2 values between 0.61 and 0.86. It should be noted that R^2 coefficient values were incrementally improved with the inclusion of both SPOT spectral indices and LiDAR estimates, increases ranging from 0.2-0.4.

A relatively recent approach to the quantification of forest structure using LiDAR has been the ‘binning’ of the normalised 3D point cloud to reduce the data volume to a single measurement (Chasmer et al., 2004; Lee et al., 2004; Popescu and Zhao 2008; Wang et al., 2008). This method classifies height ranges within the canopy as a set of volumetric pixels, or

voxels. A voxel can be conceptualised as 3D cubes with X, Y and Z dimensions that can be used to classify laser pulse counts within the voxel extent. Voxel size can be altered, for example to represent where LiDAR returns are clumped, where smaller voxels can be used to classify dense clusters.

Another measure having a direct physical basis is canopy cover, calculated as a percentage of LiDAR returns intercepted by the vegetation canopy, within a vertical height bin size (grid cell resolution) specified by the user. The amount of light penetrating the forest canopy and reaching the ground has a strong influence on understorey composition (Martinuzzi et al., 2009). Hudak et al. (2009) state that vegetative cover at various heights for example can be placed in height ‘bins’ and output as two-dimensional raster layers. Thus, the distribution of canopy height values within a bin is effectively a ‘structural signature’ analogous to the ‘spectral signature’ of a multi or hyperspectral image pixel, although describing very different vegetation properties.

Lee et al. (2004) utilise the voxel-based approach for quantifying vertical forest stand structure. Small footprint first and last return LiDAR with a nominal spatial resolution of 1m was used in a mixed species forest dominated by white cypress pine (*Callitris*), smooth barked apple (*Angophora costata*) and eucalyptus (*Eucalyptus globus*) tree species in central Queensland, Australia. The voxel matrix was intersected with a tree crown map derived from fieldwork, classifying cells as either trunk or canopy. Various attributes from each of the individually identified trees could then be extracted, such as crown dimensions, species and growth stage. A value of crown openness was calculated for each of the individual trees, based upon published records for selected species.

LiDAR has the ability to measure the vertical structure of forest stands accurately (Lim et al., 2003b; Naesset, 2002). These approaches typically focus upon the characteristics of the top canopy layer. Methods have been presented for estimating vertical canopy structure from the raw point cloud, addressing the major task of detecting the number of main canopy layers and the height range of each layer. For example, Wang et al. (2008) proposed a method using a statistical process on a normalised point cloud which was segmented into a number of tiles due to the volume of data. Canopy layers were then detected using a salient curve detection algorithm based on a height distribution probability function. The number of canopy layers for each tile and the height range of each canopy layer were the main attributes to be derived.

Individual tree detection

LiDAR sensors, typically DR, can collect data at point densities sufficient to identify individual tree crowns in open canopies. The primary hurdle in large scale applications is the separation of tree crowns in dense forest (Maltamo et al., 2004; Wang et al., 2004). In dense closed canopies tree crowns overlap, causing a model to perceive several trees as one, producing high rates of commission and omission errors in estimated single tree attributes from canopy height models. Recent developments in the computer analysis of high spatial resolution DEMs (equal or less than 1m^2 per raster cell) are leading towards the semi-automated production of forest inventory information from remote sensing, which can be divided into finding tree locations with crown size parameterisation or full crown delineation (Gougeon and Leckie, 2003; Hyypä et al., 2008). Methods already developed using high spatial resolution aerial optical imagery can be utilised for this purpose. Alternatively, in laser scanning it is possible to improve the image-based approaches by using statistical algorithms, voxelisation, and knowledge-based approaches.

There is an extensive literature concerned with the automated extraction of features from elevation data, but only a subset of these studies have focused upon natural environments (Koukoulas and Blackburn, 2005a). A number of organisations, especially the Canadian and Australian forest organisations, have driven the development of many of these tree-crown detection algorithms (Bunting and Lucas, 2006).

In research presented by Kaartinen and Hyypä (2008) a high degree of variation in the quality of published methods was found between studies with the percentage of correctly delineated trees ranging from 40-93%. Kaartinen et al. (2012) reviewed nine separate methods of individual tree detection in the literature based upon both image-based and point-cloud-based methods. Within the results of this work, many approaches were found to provide a means to delineate dominant trees in terms of location, height and crown extent. However, crown size estimates could vary significantly. The highest accuracy was achieved using a minimum curvature-based approach which was applied to the canopy height model (CHM), where the minimum value coincides with the tree top, and the maximum value coincides with the valleys surrounding the crown. This was accompanied by point-cloud-based cluster detection for suppressed trees. FW technology is expected to improve individual tree detection, especially in the case of suppressed trees, as waveform analysis can be used to provide denser point clouds within the tree crowns.

Forest gap characteristics

A number of studies have focussed upon the importance of gap openings within the forest and demonstrated the associations between the gap characteristics and ecosystem functioning and the importance of gaps in determining species composition and controlling regeneration processes (Runkle, 1982; Platt and Strong, 1989).

The work presented in Koukoulas and Blackburn (2005b) demonstrates the capability of LiDAR for the detection of forest gaps using GIS contouring procedures. Forest gap size, shape and height diversity (within the gap) parameters, in addition to tree size parameters were extracted for a deciduous tree species dominated area within the New Forest, UK. The relationships between the spatial patterns of gaps and trees were characterised using a Ripley's K-function analysis. The researches findings indicated that large gaps and large trees are randomly distributed, whilst smaller trees and smaller gaps are clustered. Significant relationships were also recorded between patterns of tree species and different size classes, as well as between patterns of trees and gaps with specific properties. The quantification of these gap characteristics alone is insufficient to provide conclusive evidence on specific processes, but Koukoulas and Blackburn (2004) state that such information can be indicative of the general status of a forest and can provide new perspectives and possibilities for further ecological research and forest monitoring activities.

2.3.2.4 Forest composition

LiDAR metrics relating to return heights and intensity, in addition to waveform metrics for FW data, can potentially be used to classify image object types. Considering the case of FW LiDAR, in theory this technique provides the amplitude and echo width in addition to a 3D point cloud (Wagner et al., 2008; Wagner et al., 2006). The amplitude provides information on the target's reflectance, and the echo width is a measure of the range variation of scatterers within the laser pulse footprint contributing to a single echo and is therefore an indicator for surface roughness (Chauve et al., 2007), although there may be issues in transferring the method to other sites.

Forest tree species classification

In many forest applications, tree species is of particular interest as an essential component of forest studies, inventories and managements. At present, species classification can be accomplished using both optical and laser data. Conventionally, tree species information is extracted from high spatial resolution aerial photographs or multi-/hyper-spectral imagery. A number of pieces of research have focused upon classifying tree species using LiDAR. Holmgren and Persson (2004) for example state that it is possible to separate coniferous and deciduous species in the Reminigstorp area, Sweden, using near infrared images. However, the coniferous species of pine and spruce have similar spectral signals. The structure and shape of the tree crowns can be extracted from airborne LiDAR data and used for the discrimination between spruce and pine species at an individual tree-level. The proportion of correctly classified trees was 95%.

Research presented by Moffet et al. (2005) explored the use of the proportion of single returns from multiple return LiDAR systems, as an important predictor for tree species classification, namely poplar box (*Eucalyptus populnea*) and cypress pine (*Callitris columellaris*) species. This research was conducted in South-East Queensland, Australia. A map of different returns overlaid upon field species data indicated it was possible to distinguish between vegetation types that produce a large proportion of single returns, compared with vegetation types that produce a lower proportion of single returns. While a clear distinction was not always obvious at the individual tree-level for the two species, due to extraneous sources of variation in the dataset, the observation was supported in general at the site-level. Sites dominated by poplar box generally exhibited a lower proportion of singular returns compared with sites dominated by cypress pine. This research also explored the use of return intensity statistics; unfortunately this was found to be less useful for classification purposes.

Most discrete return LiDAR systems also record the return intensity of the laser pulse. Thus LiDAR return intensity values are becoming increasingly used. This may aid in interpreting a particular sample pulse as belonging to either ‘ground’ or ‘crown’ classes, considering that each surface will produce differing return intensities. Intensity is the ratio of the power returned to the power emitted and is mainly a function of surface reflectivity at the emitted wavelength. It is also a function of the area of the object that returns the pulse, and the portion of the pulse remaining after previous returns (Kaasalainen et al., 2009; Kim et al.,

2009a). Intensity data are not generally calibrated for differences in receiver gains and are periodically adjusted during acquisition. Gain settings are currently proprietary, and typically are not made available to the end user. The use of intensity data is complicated because of a variety of factors affecting the reflected laser signals from the surface including range, incidence angle, bidirectional reflectance, atmospheric transmittance and attenuation, transmitted power and beam divergence (Hyyppä et al., 2008). If these factors are compensated for, intensity values can be used effectively. The majority of airborne LiDAR systems use an infrared laser, thus the return intensity of the reflection is sensitive to vegetative materials (Andersen et al., 2006). Kaasalainen et al. (2009) illustrated the potential to calibrate this return intensity data with reference targets to enable classification of land cover types, such as trees and vegetation.

Brandtberg et al. (2003) utilised a method focusing upon individual trees, creating tree crown polygons from applying an automated procedure to the LiDAR derived CHM, which is discussed in greater detail in the next section. The intensity of the near-infrared reflection from the LiDAR data acquired in leaf-off conditions were used to determine species type in a West Virginian deciduous forest, USA. LiDAR return intensity successfully discriminated leaf-off oak, maple and poplar species to a moderate to high degree of accuracy.

Vaughn et al. (2012) documents another approach to use either DR or FW LiDAR for species classification for a total of five coniferous and deciduous tree species in the Pacific North-West United States. The authors state that LiDAR data can be used to extract: crown density, crown shape, crown surface texture, and received energy from individual peaks (i.e. return intensity). These metrics are summarised in Table 2.6. A computationally intensive method of producing a 3D voxel-based region-growing segmentation algorithm for identification of individual tree crowns was instituted. Various point and voxel cluster characteristics relating to spatial arrangement and waveform statistics, such as amplitude, width, and distances between returns in a waveform were extracted and utilised as input into the classification algorithm, in addition metrics from a Fourier Transform analysis were applied to the individual waveforms. Using a Support Vector Machine (SVM) classification, Vaughn et al. (2012) achieved a classification accuracy of tree crowns for DR data alone of 79.2% overall ($\kappa = 0.74$). The incorporation of waveform information improved the overall accuracy to 85.4% ($\kappa = 0.817$) for five tree species, two were coniferous (douglas fir and western red cedar) and three were broadleaved (bigleaf maple, red alder, and black cottonwood).

Table 2.6 – Summary of tree canopy descriptor concepts

Key Concept	Description
<i>Crown density</i>	Describes leaf and branch size and arrangement and is typically measured using the proportions of the returns hitting vegetation versus those hitting the ground (Brandtberg et al., 2003; Moffiet et al., 2005).
<i>Crown shape</i>	This information is often compared using parameters of surface models fit over the top of the LiDAR point cloud (Holmgren and Persson, 2004; Reitberger et al., 2009).
<i>The distribution of return heights</i>	This metric is often described using selected percentiles of the return heights which include information about both crown density and crown shape (Korpela et al., 2012).
<i>Crown surface texture</i>	Crown surface texture refers to the roughness of the tree crown surface, and has been measured using a canopy height model (Vauhkonen et al., 2009).
<i>The measured intensity</i>	The measured intensity is affected by several physical traits such as leaf size, chemistry, and incidence angle, which in turn are effected by species type. While most researchers incorporate intensity, a number of authors have found intensity alone could be a reasonable predictor of species (Kim et al., 2009b).

2.3.2.5 Forest understorey and deadwood components

Mapping the distribution of snags and understorey shrub species across the landscape has presented a major challenge in previous research for the management of biodiversity and wildlife habitat. Characteristics associated with forest structure, such as the height of trees, presence or absence of understorey, canopy closure, tree diameter, abundance and size of dead trees, etc. are factors which may explain the presence of many species of wildlife, the functional use of the habitat and the overall diversity of wildlife species in the forest (Martinuzzi et al., 2009). Additionally, animal use of different understorey does vary. Recent studies have been able to characterise height, understorey vegetation cover and the detection of suppressed trees (Hill and Broughton, 2009; Maltamo et al., 2005). This work was accomplished through the use of canopy height thresholds, cluster analysis and visual interpretation. These studies have shown that assessments of understorey vegetation with LiDAR is less accurate under dense tree canopies due to the lower proportion of LiDAR returns reaching the lower forest strata (Maltamo et al., 2005).

Hill and Broughton (2009) showed it is possible to characterise understorey vegetation in closed deciduous forests in Monks Wood, UK, through the integration of leaf-on and leaf-off LiDAR data. Leaf-off data were acquired at a time when the understorey had begun to leaf-up, thus representing the understorey and ground better than the leaf-on data (the understorey and overstorey canopies in lowland broad leaved woodland can merge into one another). Comparisons between different species groups were made using an ancillary tree species map

and field data. The resultant understorey map had a 72% correspondence with field data on the presence/absence of suppressed trees and shrubs.

Martinuzzi et al. (2009) utilised a method to use LiDAR to derive a variety of environmental factors to explain the presence/absence of understorey vegetation and standing dead trees (snags), in addition to testing it as an input for a habitat suitability index (HSI) for bird species. HSIs have been used to assess the potential of an area to support the resource, shelter, or reproductive needs for a given wildlife species, the output of which is a map depicting suitability values across the landscape for the target species. Fieldwork measures were statistically regressed against LiDAR metrics for forest structure and topography using the random forest algorithm in the statistical package R (Hudak et al., 2008a). The understorey shrub presence/absence prediction yielded an overall accuracy of 83%. Similarly, snag classification yielded accuracies ranging between 72% and 80%. The HSI map produced had an accuracy range of between 79% and 90% depending upon the species.

Korpela et al. (2012) outlined a method of estimating the number of understorey trees in multi-layered pine stands using small-footprint FW LiDAR data. The overstorey layers often obstruct ‘wall-to-wall’ sampling of the understorey using LiDAR because of transmission losses which affect triggering probabilities and peak-amplitude (intensity) observations. The study employed a method where transmission loss compensation models were based on the returned power of each return pulse waveform and the total, employing the geometry of the pulses in relation to reference trees and overstorey intensity observations as predictors. The analysis was carried out via a custom Java program, where each pulse echo was assigned to objects which triggered it (e.g. a reference tree). The area-based LiDAR height metric of the proportion of ground returns had correlations with understorey stem density for low standing trees (height >0.3m), where the R^2 value was 0.55-0.87 ($p < 0.05$). Detailed knowledge of trees within an area is required for this method, and as such it may be impractical for cost- and temporally-efficient assessment methods.

It should be noted that a variety of environmental factors can influence the presence of coarse woody debris (CWD) and understorey shrubs in the forest and therefore have the potential to serve as predictor variables in a distribution modelling approach. An evaluation of the structure and composition of understorey vegetation found that the overstorey canopy structure, topography and land use can all influence the understorey in forests. Martinuzzi et

al. (2009) determines that understorey is denser in open forests, where more light can penetrate to the ground, highlighting the importance of estimating overstorey density.

With regard to the presence of deadwood in the overstorey and understorey, accurate estimations of alive and dead biomass in forested ecosystems is important for studies on carbon dynamics, biodiversity, habitat assessment and wildfire behaviour. Traditional methods of measuring deadwood, or CWD, are expensive and suffer from low accuracy as dead trees are often rare and their existence is clustered (Pesonen et al., 2008). LiDAR remote sensing has been used successfully to estimate live biomass, whereas studies focusing on dead biomass are rare (Kim et al., 2009b; Pesonen et al., 2008). Pesonen et al. (2008) outlined a project where estimates of downed and standing deadwood volumes were investigated in the Koli National Park, Finland. Using field data gathered from 33 plots and site fertility estimates, variables for CWD and living biomass were produced and used to create regression models with various LiDAR height and intensity metrics. The resulting accuracy was stated to be adequate for predicting downed deadwood volumes (RMSE 51.6%); however the standing deadwood volume estimates were poorer (RMSE 78.8%). A similar approach was also attempted by Kim et al. (2009) for mixed coniferous woods of the Grand Canyon National Park, USA, including the use of LiDAR intensity metrics. This study provided slightly better estimates of CWD through the study area.

Mücke et al. (2012) described research into mapping fallen dead tree stems within North-East Germany using very high density FW LiDAR acquired from a helicopter platform for leaf-on (21.8 echoes/m²) and leaf-off (16.9 echoes/m²) conditions. Under the Natura 2000 framework, the assessment of ecosystem condition is of paramount importance. One such indicator is the abundance of deadwood in forested ecosystems, which was identified as an important indicator for habitat condition (European Commission, 2012). The orientation and processing of airborne laser scanning data (OPALs) software was utilised for all processing tasks. In order to detect downed dead tree stems the work of Wagner (2006) was utilised. FW return echo width relates to small height variations of scattering elements within the footprint of the laser beam, and was considered a means of inferring surface roughness. Forest ground-level and downed stems were assumed to have smooth surfaces, whereas other vegetated elements, such as shrub vegetation, were considered to be rougher. For the whole study area, out of a total of 193 manually digitized stems, 72 stems were fully detected (37.3%), 64 partly (33.2%) and 57 were not found (29.5%).

2.3.2.6 Attribute summary

Airborne DR and FW LiDAR technologies have been applied to extracting many forest attributes, some of which are not mentioned here, such as the modelling of successional stages (Falkowski et al., 2009a; Zimble et al., 2003). In terms of metrics related to assessing forest condition, examples of LiDAR based research have provided many possible metrics and methods for many of those variables listed in Table 2.4. Research identified from the surrounding literature provide examples of LiDAR techniques to estimate forest structure for both individual tree and area-based metrics, such as the numbers of trees within an area, metrics of tree size, e.g. basal area, number of canopy layers, average stand height, and canopy dimensions. Forest compositional metrics have been estimated for the classification of small numbers of species and land-cover types. The composition of forest understorey components such as volumes of deadwood and understorey vegetation structure was also covered.

2.3.3 Fusion of optical and LiDAR data

Fusion between LiDAR and other data sources is becoming a research topic in itself. Some studies concern the simultaneous use of LiDAR and multi or hyperspectral datasets, while others consider the combination of 3D information obtained from LiDAR and photogrammetric techniques (St-Onge et al., 2008) using both aerial and satellite imagery. The combination of the 3D LiDAR and 2D spectral information is an area of significant potential. For example the expected combination of information from optical and LiDAR data will be better able to delineate tree crowns from above (Holmgren et al., 2008; Leckie et al., 2003a). Another area of research has involved the use of combining optical multispectral, LiDAR and radar data. Hyde et al. (2006) outline an approach combining such information sources statistically in order to extract forest structure information for wildlife habitat analysis. The results concluded that LiDAR was the best single sensor for estimating canopy height and biomass. With the addition of multispectral data, improvements were made in the estimation of tree structure. Hyde et al. (2006) concluded by saying that the structural metrics extracted from LiDAR combined with radar were essentially redundant.

Fine resolution multi or hyperspectral imagery and LiDAR data capture differing, yet complementary characteristics in the estimation of forest structure, often described as synergistic (Anderson et al., 2008; Coops et al., 2004). The use of each of these systems addresses some the limitations of the other. Most information on individual tree species,

health and spatial geometry of the tree crown is provided by multi or hyperspectral imagery. LiDAR however, provides data concerning tree height, 3D crown shape and texture or outline (Hyypä et al., 2004; Leckie et al., 2003a). It should be noted that LiDAR and optical data do not have to be collected at the same time (Hyypä et al., 2004). Non-coincident data capture, for example, optical imagery collected in the vegetation growth season with LiDAR data captured over the vegetation in leaf-off conditions, allow obvious advantages in winter months where lasers can penetrate the canopy of deciduous trees so that vertical structure can be better discerned (Brandtberg et al., 2003).

Only a small amount of literature has been published on the topic of fusing different sensor technologies together for forest applications (Hyypä et al., 2004). Of the published literature concerning data fusion, examples have been presented where high resolution optical data have proven to better outline tree crowns in dense forest situations when compared with LiDAR. LiDAR data however, can reduce the commission errors that occur in open stands with optical imagery, for example the application of a height filter to remove sub-canopy vegetation (Hyypä et al., 2004; Leckie et al., 2003).

The following sections identify those fusions of multi or hyperspectral and LiDAR remote sensing case-studies which extract or estimate those forest condition indicator metrics stated in Table 2.4.

2.3.3.1 Forest structure

Lucas et al. (2008b) utilised a combination of CASI-2 multispectral data at a nominal spatial resolution of 1m and covering the spectral range of visible to near-infrared, and DR LiDAR data acquired using an Optech 1020 scanner with a sample density of a point every ~1m. The aim of this research was to develop a method to extract estimates of biomass for individual trees for mixed species forests in Queensland, Australia. Individual trees were identified using a combination of the tree crowns delineated automatically through segmentation using the CASI data and stems located using a LiDAR height scaled canopy openness index (HSCOI). Tree species information was then extracted for each of the identified tree crown objects using the multispectral data, unless it was a suppressed tree. The component for biomass of individual trees was estimated using LiDAR-derived height and stem diameter as input to species-specific allometric equations. These estimates corresponded to plot-based estimates with an R^2 value of 0.56. Additionally, a second approach utilised a jack-knife

linear regression using LiDAR-derived heights and crown cover at the plot scale and produced more robust estimates of biomass ($R^2=0.90$). A number of issues were highlighted in terms of over/under-estimations from the LiDAR data dependent on forest cover species type and stem density due to the complexity of the woodland and to the sometimes poor correspondence with objects generated through segmentation.

Popescu et al. (2004) proposed a fusion of small-footprint LiDAR and multispectral data to estimate timber volume and biomass at the plot-level in deciduous and coniferous forests in Virginia, USA. Individual tree heights and crown diameters were estimated using ITC algorithms. An assumption was made that there was a relationship between the height of the tree and its crown size. A regression analysis was then performed to relate field measures of tree DBH and height against those retrieved from the remote sensing data. Regression models and cross validation were then used to estimate inventory data including volume (R^2 deciduous 0.39; coniferous 0.83) and biomass (R^2 deciduous 0.32; coniferous 0.82).

A number of studies have focused upon combining LiDAR and spectral data for the purpose of modelling fire behaviour. In order to satisfy the criteria laid out for fire monitoring applications, factors such as canopy height, vegetation type, dead and live fuel load, and percent canopy cover must be estimated. Mutlu et al. (2008) utilised raster products from multispectral Quickbird satellite image bands and multiple LiDAR height bins and canopy height for characterising forest fuels, for a study site in eastern Texas, USA. A combination of Principle Components Analysis and minimum noise fraction was used to remove the least relevant raster products in an attempt to improve supervised classification accuracy. This resulted in an accuracy of 90.1% with the fusion of airborne LiDAR and satellite Quickbird imagery.

2.3.3.2 Forest composition

LiDAR data can be used to predict the species of individual trees, at least in boreal forests with relatively few tree species (Hyypä et al., 2008). Such predictions can be improved using a fusion of LiDAR and optical imagery. However, dense LiDAR data have been shown to produce accurate tree-level classification, even without optical imagery (Hyypä et al., 2008; Reitberger et al., 2008).

Hill and Thompson (2005) outline a method utilising both airborne hyperspectral data acquired using the HyMap sensor, at 4m spatial resolution covering the spectral range from visible to shortwave infrared (0.437-2.486 μ m), and discrete return LiDAR data, from an Optech ALTM 1210, with an average point density of one point for every 4.83m². Their research investigates the ability of the combined dataset to generate unique thematic classes based upon the tree and shrub species composition and vegetation structure, for a site in Cambridgeshire, UK. Classification of the ten various thematic classes was based on the National Vegetation Classification (NVC) scheme for woodlands and scrub. This was achieved using principle components analysis and LiDAR-derived canopy height models. Automated segmentation algorithms were then applied to identify spatial groupings with similar characteristics. It should be noted that this method does not provide information on ground flora composition.

Dalponte et al. (2008) proposed an alternative method in fusing high spatial resolution airborne hyperspectral and LiDAR remote sensing for classification of tree species in complex environments, in this case a nature reserve in the Po Plain, Italy. A software system was produced to provide inputs from hyperspectral image bands and LiDAR-derived canopy height and intensity raster layers for two tree species classification approaches, these were Gaussian maximum likelihood and Support Vector Machines. A total of 19 tree species classes were extracted and assessed against field validation data. The combination of the two data sources resulted in increased classification accuracy, over using one data source only, particularly in relation to the discrimination of very similar species. The kappa accuracies obtained with different classifiers were as high as 0.89 when incorporating hyperspectral and LiDAR layers.

Simonson et al. (2013) addressed the potential for remote sensing with regard to Natura 2000 habitat monitoring objectives. The potential benefits include the cost-effective production of habitat distribution mapping; in addition to providing biophysical indicators of functioning relevant to favourable conservation status such as LAI and vegetation fractional cover amongst others. The research tests the complementarity of multispectral and DR LiDAR in providing a robust indicator of conservation status, in this case an estimate of species richness. A raster-based analysis was conducted upon LiDAR data, where raster layers were derived for measures of image texture and the maximum, minimum, standard deviation of the first return heights and mean of last return heights. Dimension reduction was performed upon

the multispectral data in the form of Principle Components Analysis (PCA). A supervised maximum-likelihood classification was then performed upon the combined dataset creating a total of 11 land-cover classes relating to semi-natural forest, plantation forest, agricultural land, rocky surfaces and urban environments. A land cover classification accuracy of 70% was achieved using the combined dataset. Then utilising known relationships ($r = 0.75$, $p = 0.001$) between mean vegetative height, overstorey and understorey species richness, and spatial aggregation calculations using FRAGSTATS, were combined to provide a proxy indicator was created for habitat condition. The result of this calculation was a map with three classes relating to high, medium and low condition.

2.3.3.3 Summary of attributes

Example case studies concerning the combination of hyperspectral and LiDAR datasets are limited, especially for the application area of forest research. Of those relevant example case studies mentioned here for forest condition assessment, many were concerned with the enhancement of tree species classification. Other examples included the research of Lucas et al. (2008b) and Popescu et al. (2004) into the extraction of individual tree parameters and estimates of biomass. Of particular note was the research performed by Simonson et al. (2013) in order to provide forest biophysical monitoring information for a conservation initiative.

2.4 Conclusions

The traditional methods of assessing woodland condition require the collection of data manually through fieldwork. While this approach is often highly accurate, it is labour intensive, in addition to having a number of issues related to representativeness over wider areas. Remote sensing technologies are recognised as powerful tools for ecological investigation at wider spatial scales. They are seen as providing the only realistic, cost-effective means of acquiring continuous data over large areas (Ker and Ostrovsky, 2003). To date however remote sensing has seen limited uptake in the discipline of ecology. A greater integration of remote sensing and ecology disciplines should be very valuable beyond the traditional landcover mapping and pattern description (Newton et al., 2009b).

The analysis of airborne multi and hyperspectral data has well-developed and understood techniques for vegetation mapping, following the development of new technologies and techniques over the past 30 years for forest ecosystem management applications. This

development has produced mature techniques for the extraction of metrics such as the classification of species and the extraction of biophysical variables such as LAI, biomass, or chlorophyll content.

Analysis of DR LiDAR is quickly gaining prominence in natural resource research and management due to its inherent ability to represent complex 3D structures and ground surfaces with very high precision (Evans et al., 2009). These techniques are able to estimate metrics related to canopy heights, the ground surface underneath the canopy, canopy 3D structure, the detection of vegetation layers, and the estimation of attributes such as basal area, stem volume and biomass, amongst others. The ability to detect individual trees from LiDAR has seen recent activity (Kaartinen et al., 2012).

There have been relatively few studies involving small-footprint FW LiDAR, especially so for ecological analysis, which is a method which offers a number of advantages over the conventional DR systems. One of the main advantages is in theory a higher number of return points per pulse than conventional DR systems; this is due to the removal of issues related to dead-zones inherent in deriving multiple returns from DR pulses. FW data offers greater potential for detecting a greater number of surface interactions from the entire returned waveform. This is in addition to added metrics related to the returned waveform, such as pulse-width. Studies using FW data have been successful in the detection of understorey components (Korpela et al., 2012) and the characterisation of 3D canopy structure (Lindberg et al., 2012) for example.

There are a small, but growing number of studies concerned with the integration of remote sensing techniques. Fine resolution multi or hyperspectral imagery and LiDAR data capture differing, yet complementary characteristics in the estimation of forest structure (Anderson et al., 2008; Coops et al., 2004). Application areas include the delineation and species classification of individual trees (Lucas et al., 2008b), the classification of landcover (Hill and Thompson, 2005), and the modelling of fire behaviour (Mutlu et al., 2008).

Very few projects have attempted to assess woodland condition using remote sensing. A number of recent studies have attempted to map some of the key components of condition, for example the volume of standing and fallen deadwood (Mücke et al., 2012), or through assigning a basic three-tier classification through the identification of species richness,

vegetation height and patch connectivity (Simonson et al., 2013). These studies have attempted to assess woodland condition for one potential component or at a relatively coarse scale with a basic classification.

As identified previously, woodland condition is related to a complicated arrangement of compositional and structural characteristics which is typically assessed through manual fieldwork at a small scale and is then extrapolated up to the required scale. This research project attempts to address some of the gaps in research by exploring the contribution that the combination of hyperspectral and LiDAR data can make to assessing woodland condition using a suite of potential remote sensing derived indicators within a UK context. This will comprise the extraction of remote sensing metrics at the individual tree and field-plot scales to estimate field metrics over larger areas for a site within the New Forest, UK. In addition FW LiDAR holds great promise for ecological studies, an area which has yet to be fully explored. Both DR and FW LiDAR data will be analysed and compared in their relative abilities to predict woodland attributes.

Chapter 3 – Field site and data acquisition

3.1 Study site

The study site is located within the New Forest in southern England. The New Forest lies to the west of Southampton, in south-west Hampshire and covers 37,677 hectares (Forestry-Commission 2011). The New Forest is recognised as being of international importance to nature conservation for its extensive tracts of semi-natural vegetation and ecologically valuable habitats, in addition to a large number of plant and animal species. The New Forest encompasses lowland heath, valley mire and ancient woodlands. Its extensive areas of semi-natural habitats, occurring in a complex mosaic, justify its inclusion amongst the most valuable areas for biodiversity conservation in Western Europe. It is mostly Crown property and managed by the Forestry Commission, however one quarter of the areas consists of farmland and settlements. Over centuries of human management and livestock grazing the New Forest has become a combination of heathland, ancient woodland, plantation woodland, wetlands and grassy plains, with many inclosures of both coniferous and deciduous woodland.

The New Forest lies on a series of gravel terraces overlying sedimentary sands and clays of Tertiary age, located within the Hampshire Basin. As noted by Tubbs (2001) the New Forest as an ecological system has developed under the influence of large, free-roaming herbivores, including deer as well as livestock. The present character of the New Forest is therefore strongly dependent on its history as a medieval hunting forest, and the survival of the traditional ‘commoning’ system. The crownlands include inclosures, designated for silviculture and unenclosed land over which commoners’ rights prevail. The unenclosed land is referred to by Tubbs (2001) as the largest area of semi-natural vegetation in lowland-Britain. The total approximate area accounts for 20,000 ha; within that 3,700 ha is oak, beech, and holly woodland.

The New Forest National Park was designated in 2005 and extends over 57,100 ha (Chatters and Read, 2006). The conservation value of the National Park is reflected in a variety of designations, with some twenty SSSIs, six Natura 2000 sites and two Ramsar Convention sites which intersect the Park’s boundaries. Woodlands classified as ‘ancient and ornamental’ are estimated to occupy 3,671 ha (Tubbs, 2001). Many of the largest ancient and ornamental woodlands are distributed in a belt centred around Lyndhurst (Peterkin et al., 1996). The

same general pattern has persisted for at least 200 years, although during that time some ancient woodlands have been incorporated within silvicultural inclosures and replaced by plantations. Ancient woodlands are dominated by beech (*Fagus sylvatica*), oak (*Quercus robur* and *Quercus petraea*), birch (*Betula pubescens* and *Betula pendula*) and holly (*Ilex* spp). Typically beech and oak dominate the canopy, with birch occurring on the edges of woodland blocks and holly occupying the understorey.

Many of the ancient woodlands are classified as such because they have been in existence for at least 400 years, some of which can be considered direct descendants of the original 'wildwood'. While the structure and composition of the woodlands have been greatly influenced by a long history of human intervention, and in particular the activities of livestock and deer, which greatly modify forest understorey, the presence of large trees and large quantities of deadwood create a structure which is said to resemble wild ancient woodland. The New Forest inclosure comprises woodland communities which are not subject to livestock grazing until most trees are past browsing height. The inclosures are of relatively recent plantations on former heathland or ancient woodland stands, fenced off to the commoners' animals, but accessible by deer.

The unenclosed forests are permanently open to grazing by the ponies and cattle of the commoners (Forestry Commission 2011). In recent years some 6,000-7,400 ponies, cattle, donkeys, pigs and sheep have been kept in the forest by about 550 commoners. Heavy grazing and browsing by ungulates is of concern within areas of the New Forest, this is because of the impact it has upon the regeneration of tree species according to the long term study into Denny Wood of Mountford et al. (1999).

This study is focused upon an approximately 22 km² area including the Frame Heath and Hawkhill, New Copse, Parkhill and Denny inclosures, in addition to Tantany, Denny and Frame Woods which are unenclosed. The study site bounding box corresponds to north-west corner 437455E: 107543N, and the south-east corner 432529E: 100857N (see Figure 3.1). A number of photographs are presented in Figure 3.2 to exemplify the various forest structural types and environments within the New Forest.

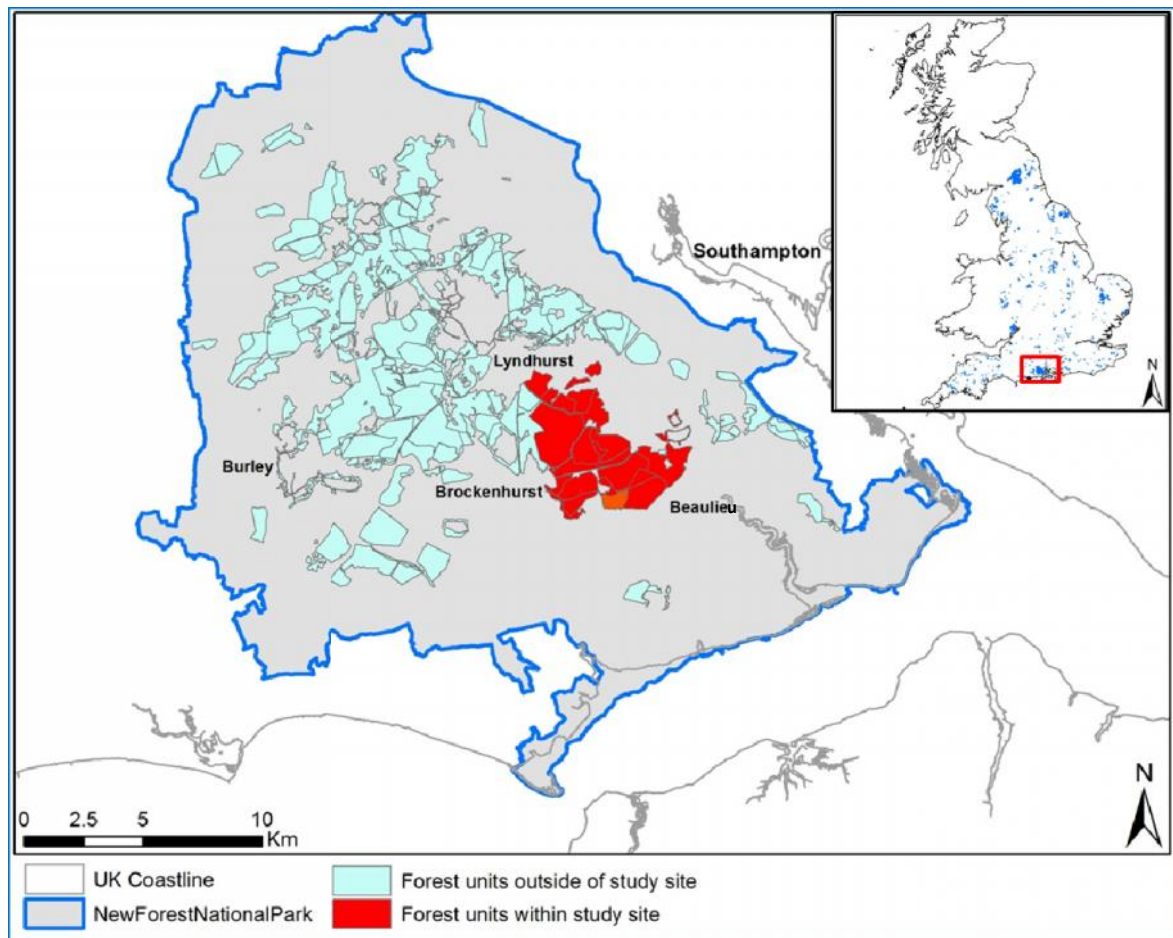


Figure 3.1- Study site locations within the New Forest

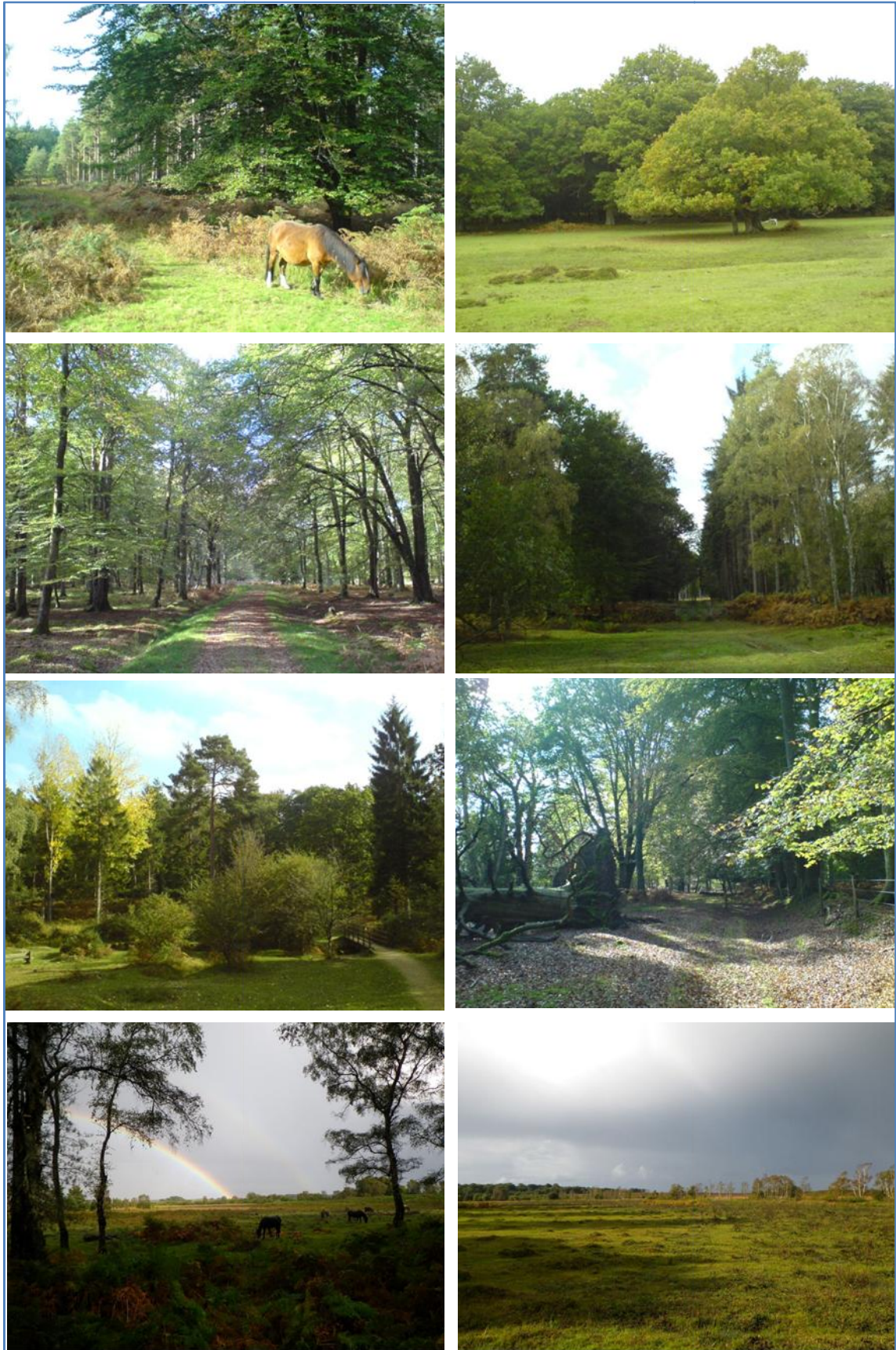


Figure 3.2 – Various photographs taken during fieldwork to exemplify the various forest structural types and environments surrounding the forest.

3.1.1 Field-plot site locations visited in 2010

The initial study sites focused upon an area including the Frame Heath and Hawkhill inclosures, in addition to Tantany and Frame Woods (see Figure 3.3). These woodlands contain several types of semi-natural and plantation coniferous and deciduous woodlands.

- Frame-Heath inclosure (~175ha) (5 plots): was first enclosed in 1852. The site contains conifer stands of modest size, primarily containing corsican pine (*Pinus nigra* *maritima*), scots pine (*Pinus sylvestris*), douglas fir (*Pseudotsuga menziesii*) and norway spruce (*Picea abies*). The inclosure also contains a considerable number of mature broadleaved trees, primarily oak (*Quercus* spp.).
- Hawkhill inclosure (~190ha) (11 plots): was first enclosed in 1870. This inclosure is principally coniferous, comprising corsican pine, scots pine, douglas fir, and norway spruce. The inclosure contains a small proportion of ancient and ornamental woodlands. It should be noted that remnants of buildings remain in the south of the inclosure which potentially date back to the Second World War.
- Frame Wood (~60ha) (2 plots): is dominated by oaks (*Quercus robur* and *Quercus petraea*) and beech (*Fagus sylvatica*) with planting beginning in the 17th century (Koukoulas and Blackburn 2004; Newton et al. 2010). The area also includes holly (*Ilex aquifolium*) pasture. There was selective felling during the 18th and 19th centuries, seeing the gradual removal of older trees. Thus, the site is best described as semi-natural.
- Tantany Wood (~200ha) (3 plots): is similarly dominated by old beech, oak and holly (Newton et al. 2010). Occasional tree blow-downs have been reported in this area.

This array of forest types within close proximity of each other presents a wide range of forest structural variables of interest in this research project, such as canopy gaps and the presence of deadwood.

Both Frame-Heath and Hawkhill inclosures are worked for timber harvesting operations, the most recent of which was conducted between 2010 and 2011, at the time of writing. Frame Wood has been the subject of a number of previous remote sensing studies to determine ecological characteristics performed as part of airborne remote sensing campaigns for the UK (Koukoulas and Blackburn 2004).

It should be noted that the field plots enumerated at these sites were used to supply data for establishing relationships between field structural and compositional metrics with those derived from remote sensing datasets.

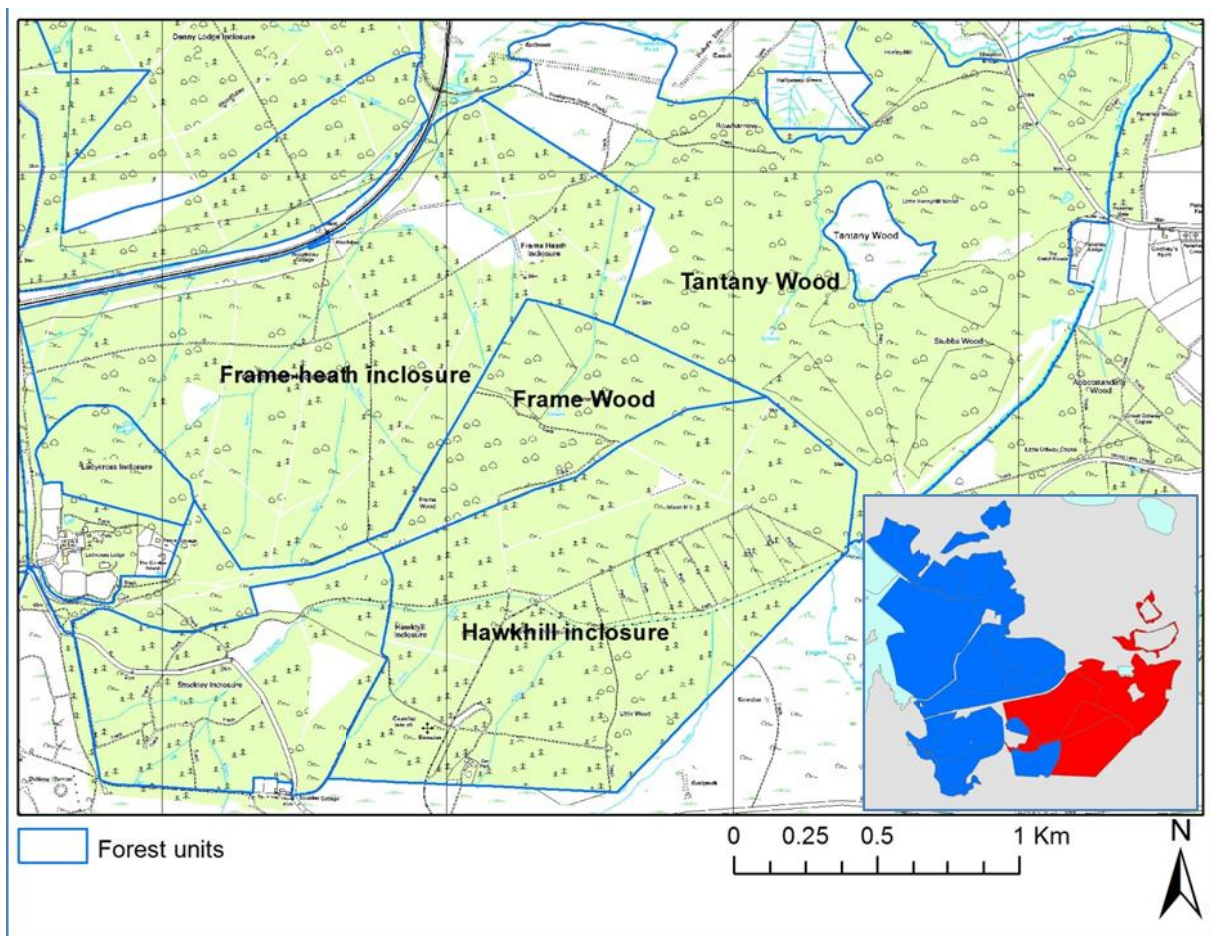


Figure 3.3 – Study site location within the New Forest. The small map in the bottom-right corner represents the total field site area; inclosures with field plot sites are indicated in red. Base Map layer is © Crown Copyright/database right 2010. An Ordnance Survey/EDINA supplied service.

3.1.2 Field-plot site locations visited in 2012

Field work designed to provide validation data was conducted within the New Copse, Denny Lodge and Denny inclosures in addition to Denny Wood between June and July 2012. Again these woodlands contain several types of semi-natural and plantation coniferous and deciduous forests (see Figure 3.4).

- New Copse inclosure (~120ha) (2 plots): first enclosed 1808. Some conifers, but a great many mature oak and beech trees, including a small number that probably pre-date enclosure.
- Parkhill inclosure (~375ha) (6 plots): first enclosed, in part, in 1751. This is an enormous inclosure that contains a mixture of broad-leaved trees and conifers.
- Denny inclosure (~125ha) (6 plots): first enclosed, in part, in 1750. Contains broad-leaved trees, mainly oak.
- Frame-Heath inclosure (~176ha) (1 plot): see above.
- Frame Wood (~60ha) (1 plot): see above.
- Denny Wood (~110ha) (1 plot): this woodland is an ancient, mixed deciduous woodland dominated by beech and pedunculate oak whose canopy trees range in age from approximately 70 years to over 300 years (Mountford et al., 1999). In addition holly is present within the understorey.
- Stockley inclosure (~33ha) (1 plot): contains a mix of semi-ancient woodlands dominated by oak and sweet chestnut (*Castanea sativa*) in addition to plantation woodland containing japanese larch (*Larix kaempferi*) western hemlock (*Tsuga heterophylla*), douglas fir and scots pine.
- Stubby-copse inclosure (~45ha) (2 plots): contains plantation woodland dominated by beech, oak, douglas fir and scots pine.

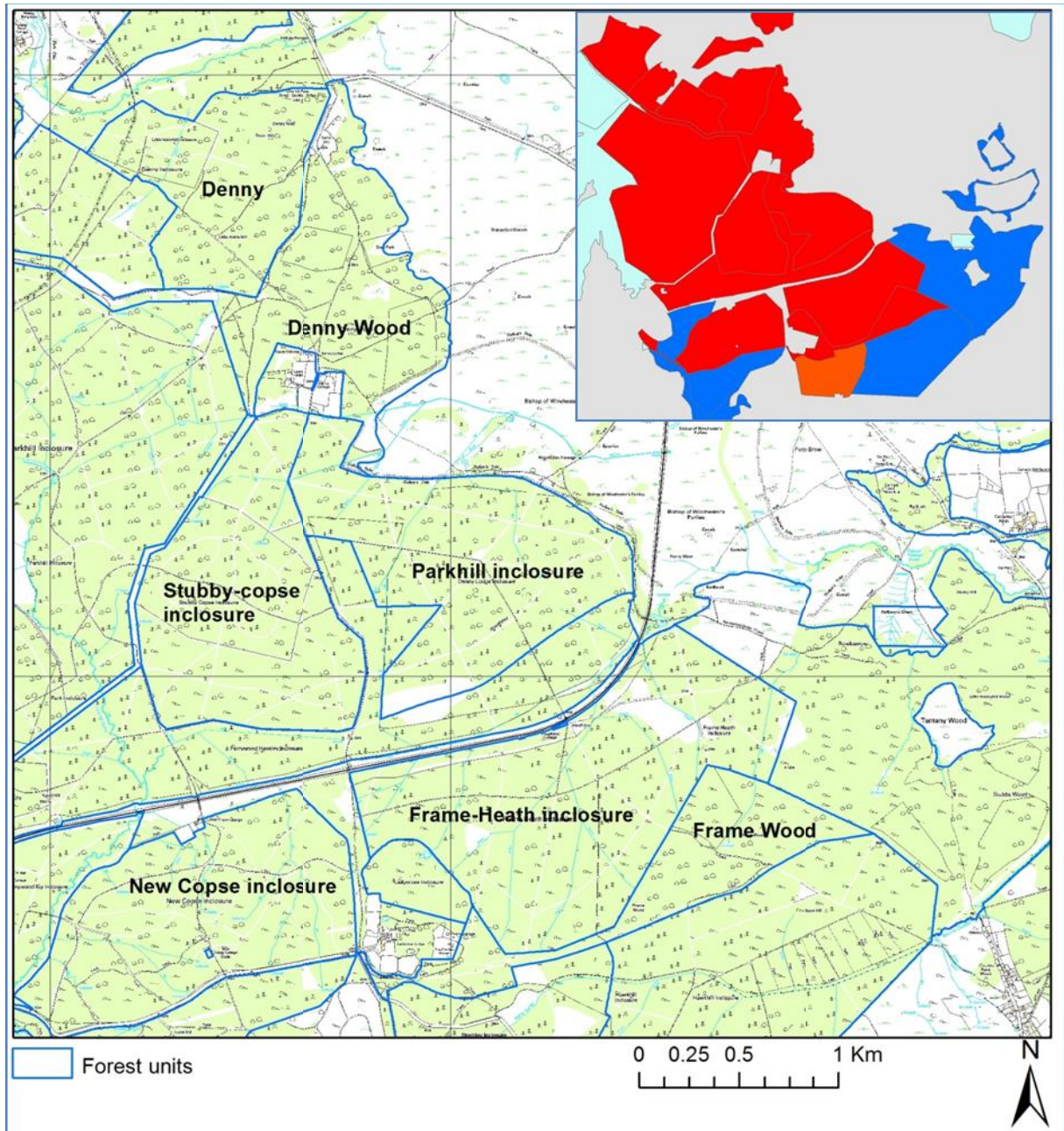


Figure 3.4–Study site location within the New Forest. The small map in the top-right corner represents the total field site area; inclosures with field plot sites are indicated in red. Base Map layer is © Crown Copyright/database right 2010. An Ordnance Survey/EDINA supplied service.

3.2 Data sources

The following section summarises each of the datasets available to and collected during the research project. These are:

1. published data from the Forestry Commission inventory;
2. fieldwork data collection;
3. remote sensing datasets (multi-/hyper-spectral and LiDAR).

3.2.1 Forestry Commission inventory data

Existing forest inventory information from the Forestry Commission's (FC) National Woodland Survey carried out between 2009 and the present was made available for all New Forest inclosures. The FC data was originally created using a base map derived from aerial photography data, with attributes populated through detailed field survey. This field survey was conducted using 1 ha sample plots, the locations of which were selected by a random process. Details such as the area, distribution, composition and condition of woodlands, together with information about the species, ages and sizes of trees were included. This dataset was compiled for both forest management and research activities (Forestry Commission, 2011).

Compartment and sub-compartment boundaries within the inclosures were provided in ESRI shape file format. Each shape file contained data giving information for each compartment and sub-compartment. Data was available at the sub-compartment scale, which included sub-compartment size (m²), sub-compartment type, plantation year, yield class, dominant-, secondary-, and tertiary-species type, and cultivation status.

3.2.2 Field data capture

The aim of the field data capture was primarily to select sites which sampled as many forest structural types as possible for an initial selection of field plots in order to provide training data for predicting field plot-level metrics over wider areas. Subsequent field data capture, in 2012, focused upon collecting data for validation of those predictive models.

The following sections define how field plot locations were initially determined and what measurements were performed within each field plot extent.

3.2.2.1 Number of field plot samples

A total of 21 field plots were enumerated in the summer of 2010. This number of sites was determined by logistical reasons, including available time, travel costs, and Forestry Commission site access restrictions due to active forestry operations. Each field site took approximately six hours to enumerate due to the large number of metrics recorded (50 in total), in addition to travel time (on foot) to and from each field plot location.

Subsequent field data capture in summer 2012 focused upon collecting data for the validation of predictive models developed using the 2010 field data. A total of 20 validation field plots were enumerated in 2012, in order to provide a similar population size to that of the initial training dataset.

3.2.2.2 Field data plot size

Guidelines from the Forestry Commission recommend a circle or square sample plot area of between 0.005 and 1 ha, which is dependent upon the objective of the survey (Matthew and Mackie, 2006). A field plot size of 50x50m was recommended by Cantarello and Newton (2008) for measuring metrics for Natura 2000 assessment with a single sub plot of 10x10m within for detailed assessment of variables with large populations and small size (e.g. counting of seedlings) due to its intensive nature. Lee and Lucas (2007) used a square field plot sample size of 50x50m with five 10x10 sub-plots within for the sampling of forest attributes in order to produce estimates from remote sensing data, whilst Naesset (2002) and Naesset and Økland (2002) utilised a somewhat smaller circular field plot size of 200m². These plot sizes can be used for extracting or estimating both area-based and individual tree metrics.

A pilot field data collection day was undertaken in 2010 to test the feasibility of recording forest information for the metrics identified in section 3.2.2.4, at location 1 in Figure 3.5. Plot sizes of 60x60m, 50x50m, 30x30m and 25x25m were assessed. Three sample metrics were selected to compare the ability of the plot sizes to capture trends, these metrics were (i) mean DBH, (ii) standard deviation of DBH and (iii) tree species number. The assumption was made that other metrics would follow the overall trends made by these three examples. Plot location 1 was composed of a mix of plantation conifers, scots pine (*Pinussylvestris*) and corsican pine (*Pinusnigra*) with small numbers of silver birch (*Betulapendula*) and beech (*Fagus*) tree species. Figure 3.6 illustrates the summary metrics recorded at all four field plot

sizes. The four plot sizes record similar area-based data overall, with only small differences observed between mean and standard deviation of DBH. Species number detected decreased with plot area size which was related site conditions. Smaller samples contained less information, where the 25x25m size was the poorest.

In addition to concerns over the time necessary to enumerate the data in the field, a field-plot sample size of 30x30m, with a single sub plot size of 10x10m, was selected as a compromise between the sample-sizes tested here and utilised in the literature. The decision of sample size was also linked to the ability of the possible use of satellite datasets in future work, for example the resolution was comparable with multispectral satellite datasets, such as Landsat ETM+.

3.2.2.3 Field data collection locations

In order to determine the Forestry Commission sub-compartments in which to locate a field plot site, a basic remote sensing analysis was carried out in order to delineate areas which corresponded to a range of different potential forest structural types. The authors Kalacska et al. (2004) and Manes et al (2010) state that Normalised Difference Vegetation Index (NDVI) values and their variability correlate with different forest canopy structural types. Thus, field plot locations were selected across a range of NDVI values.

An NDVI image was calculated from airborne CASI-2 multispectral imagery acquired in 2007 for the New Forest study site in ENVI 4.7 (ITT Visual Information Solutions) image analysis software. The details of this dataset are presented in section 3.2.3. All non-forest areas were masked and removed based on an NDVI threshold discriminating forest and non-forest. The NDVI image was then subset to produce images of deciduous or coniferous areas based on Forestry Commission mapping. This allowed an equal number of sample sites for deciduous and coniferous woodland types to be identified, thereby reducing potential bias towards one broad structural type.

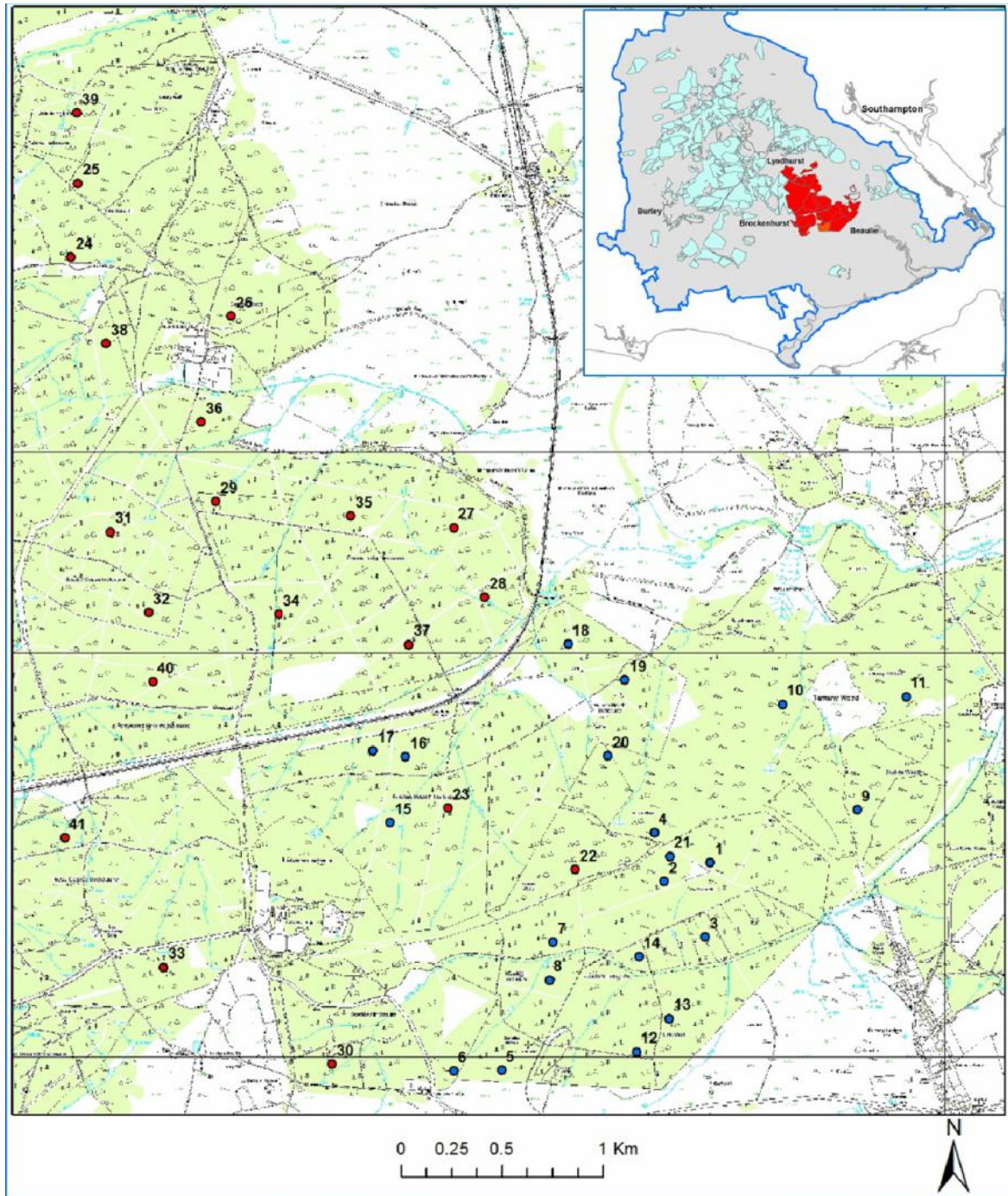


Figure 3.5 – Field plot locations. Blue points represent those plots surveyed in 2010 whilst red indicate the plots surveyed in 2012. Base Map layer is © Crown Copyright/database right 2010. An Ordnance Survey/EDINA supplied service.

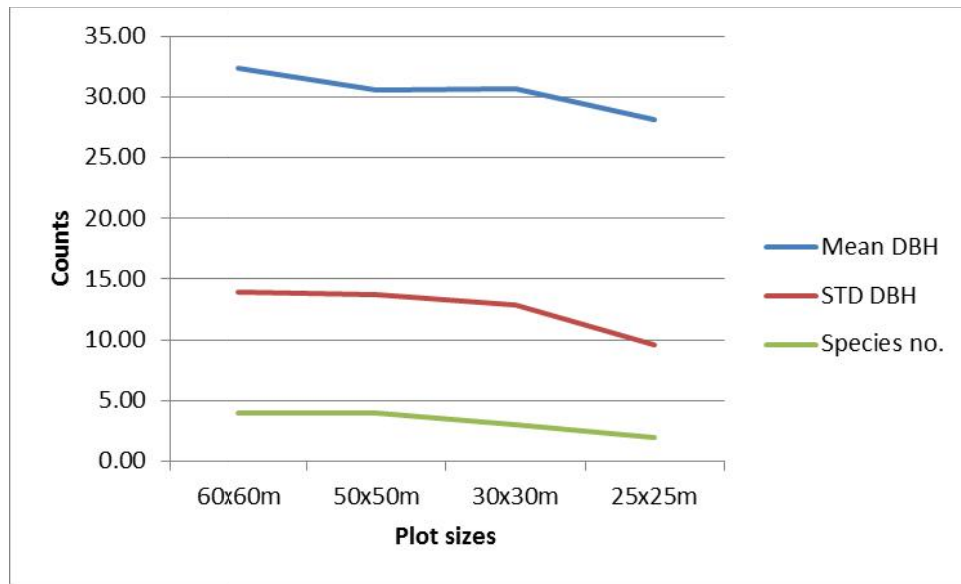


Figure 3.6 – A comparison of field plot summary statistics between four sample plot sizes.

Eight NDVI classes were created manually for the two images. The classes covered an equal NDVI range which occupied the values in-between 0.4 to 0.8 (i.e. the range of values for green vegetation in the imagery). A stratified random sampling technique was then applied to both of the classified images where an equal number of sample points were produced for each class-strata. An arbitrary minimum distance was enforced between sample points (100m) in addition to a minimum distance from the perimeter, or of no-data value (100m). The Forestry Commission sub-compartments in which a sample point was located were identified and visited within the field. Attempts were then made to locate a position as close as possible to the sample point coordinates in which to establish a field plot.

The 21 training data plots were visited during the months of June to September 2010, whilst an additional 20 plots were enumerated to provide validation data during June to October in 2012. Plot locations are presented in Figure 3.5. The plot size was established as 30m x 30m, with a 10m x 10m subplot located in the south-west corner, illustrated in Figure 3.7. To locate the extents of each of the field plots accurately, in order to best line up the data with remote sensing datasets, the Forestry Commission sub-compartment was located in the field, and the approximate coordinates identified from the generated sample point. A differential GPS (dGPS) system, the Leica GPS 500 (Leica Geosystems, part of Hexagon Group), was set up in a suitable clearing or outside of the forest in order to ensure signal errors caused by the forest canopy was reduced (Campbell and Wynne, 2011). The dGPS unit was employed to determine accurate reference points as benchmarks for locating a Sokkia 6F total station

(SOKKIA TOPCON Co. Ltd.). The total station was used to traverse from these reference points to each of the plot corners beneath the canopy, recording each in British National Grid (BNG) coordinates.

Post-processing position corrections were applied to the dGPS coordinate information using UK Ordnance Survey RINEX data using the Leica Geo Office software. Overall this provided a positional accuracy for dGPS positions of $\leq 0.03\text{m}$ overall horizontal accuracy. The appropriate coordinate corrections were then applied to total station measurements. When conducting the surveying work with the total stations, reference points were set up along the traverse, and sighted before and after the total station was relocated to provide a measure of accuracy. Through back-sighting the horizontal error was calculated as $\leq 8.11\text{cm}$ overall.

3.2.2.4 Field data measurements

There is little consensus in the literature concerning the best indicators to use for assessment of condition at the individual tree and field plot-area scales. Those used were based on the above review of the available information (see section 2.2) and are listed in Table 3.1 for structural variables, Table 3.2 for regeneration and deadwood variables and Table 3.3 for disturbance and soil variables.

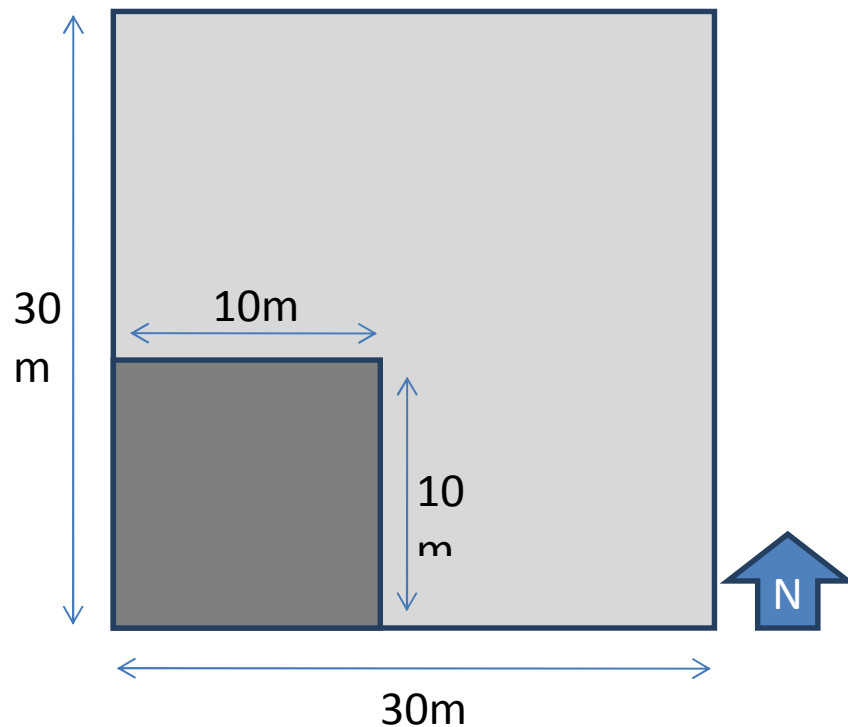


Figure 3.7 – Field plot design. 30x30m north orientated plot, with 10x10m subplot in south-west corner.

Table 3.1 – Forest structure variables recorded in the field

Metric:	Measurement type:	Plot area:
LAI Ring 4	Hemispherical camera	16 samples per plot
LAI Ring 5	Hemispherical camera	16 samples per plot
Canopy openness	Hemispherical camera	16 samples per plot
Estimate of canopy layers	Visual assessment	Whole plot
Canopy density (%)	Visual assessment	Whole plot
X,Y coordinates of each tree	Combined dGPS and total station	Whole plot
Tree species	Visual assessment	Whole plot
Girth (m) (circumference at 1.3m height)	Measurement tape	Whole plot
Tree height (m)	Combined measurement tape and clinometer	Whole plot
Height of first live branch (m) (canopy base)	Combined measurement tape and clinometer	Whole plot
North-South horizontal crown extent (m)	Measurement tape	Whole plot
East-West horizontal crown extent (m)	Measurement tape	Whole plot
Note if the tree has been 'pruned'	Visual assessment	Whole plot
Assessment of canopy condition: [Open (1) - Closed (4)]	Visual assessment	Whole plot
Number of saplings	Visual assessment	Whole plot
Number of seedlings	Visual assessment	Sub-plot

Table 3.2 – Ground vegetation and deadwood variables recorded in the field

Metric:	Measurement type:	Plot area:
Ground vegetation		
Species of shrub vegetation (if >1m height)	Visual assessment	Whole plot
Height of shrub individual (m) (if >1m height)	Combined measurement tape and clinometer	Whole plot
Horizontal extent of shrub individual (m) (if >1m height)	Measurement tape	Whole plot
Vascular plants percentage cover (%)	Visual assessment	Whole plot
Estimate of percentage bare soil (%)	Visual assessment	Whole plot
Estimate of percentage cover of moss on trees (%)	Visual assessment	Whole plot
Estimate of percentage cover of moss on ground (%)	Visual assessment	Whole plot
Deadwood		
Height of snags (m)	Combined measurement tape and clinometer	Whole plot
Girth of snag (m)	Measurement tape	Whole plot
Snag decay class: [Light (1) - Heavy (3)]	Visual assessment	Whole plot
Fallen tree length (m)	Measurement tape	Sub-plot
Fallen tree girth (m)	Measurement tape	Sub-plot
Fallen tree decay class: [Light (1) - Heavy (3)]	Visual assessment	Sub-plot
Fallen branch length (m)	Measurement tape	Sub-plot
Fallen branch girth (m)	Measurement tape	Sub-plot
Fallen branch decay class: [Light (1) - Heavy (3)]	Visual assessment	Sub-plot

Table 3.3 – Disturbance and soil variables recorded in the field

Metric:	Measurement type:	Plot area:
Disturbance		
Evidence of bark stripping (%)	Visual assessment	Whole plot
Obvious browse line on trees: [Light (1) - Heavy (4)]	Visual assessment	Whole plot
Height of browse line (m) if different from crown base	Measurement tape	Whole plot
Evidence of trampling of ground flora	Visual assessment	Whole plot
Tree seedlings and saplings projecting above ground vegetation height (y/n)	Visual assessment	Sub-plot
Evidence of browsing of tree seedling or sapling shoots (y/n)	Visual assessment	Sub-plot
Dung from animals: [Abundant (1) - None (4)]	Visual assessment	Whole plot
Evidence of human activity (e.g. vehicle tracks in site)	Visual assessment	Whole plot
Direct evidence of human activity (e.g. presence of rubbish)	Visual assessment	Whole plot
Soil attributes		
Soil pH	Tecpel PH-707 meter	4 plot corners
Soil moisture	Tecpel PH-707 meter	4 plot corners

Within each plot, the number and spatial location of each tree stem was recorded. Girth at Breast Height (GBH) of each tree >30cm GBH (i.e. approximately >10cm diameter) was measured to the nearest centimetre using diameter tape at a height from the base of 1.3m in accordance with Forestry Commission guidelines, as in Matthews and Mackie (2006). GBH can be used to determine Diameter at Breast Height (DBH) assuming the tree stem cross-section is that of a circle. Tree height and height of the first live branch (i.e. height of the living crown) was estimated for each tree (DBH>10cm) using a clinometer to measure angles from the surveyor and a measuring tape for determining distance from the tree base. Heights were derived through trigonometry (Matthews and Mackie, 2006). Tree crown horizontal extent was measured by projecting the crown onto ground-level, where the north-to-south and east-to-west extent was measured with tape (Suárez, 2004). A rough estimate of each tree's canopy condition was then undertaken according to 4 classes: (i) very sick/dying; (ii) more sick than healthy; (iii) more healthy than sick; and (iv) healthy, with little sign of deadwood and lots of green leaves (Hendry et al., 2002). The number of canopy layers was estimated at the plot-level for all species present.

The total number of saplings (and number of native saplings) were counted, with saplings being defined as trees species with DBH<10cm and >1.3m in height (Cantarello and Newton 2008). The total number of tree seedlings (and native tree seedlings) were recorded within the

10x10m sub-plot, with seedlings defined as individuals <1.3m in height. The number of ground flora species and their percentage cover was recorded within the subplot; this included vascular plants and bryophytes.

Downed deadwood (DDW) was defined as deadwood logs of at least 10cm diameter lying on the ground (Spies et al., 1988). Measurements were made in the 10x10m sub-plot only (Cantarello and Newton, 2008). Length and girth around the maximum and minimum diameters of the log were recorded. Length and girth of fallen branches >2cm in diameter were also recorded for the sub-plot. To assess DDW decay class, logs and branches were divided into three decay classes according to the following criteria: (i) logs with a low decay state, no surface breakdown, bark still intact, wood structure firm; (ii) logs with a moderate decay state, with some surface breakdown, wood structure weaker but bole mostly sound; and (iii) logs with high decay state, extensive surface breakdown, bark mostly absent, bole with no sound wood present and colonised with vegetation (Cantarello and Newton, 2008).

Snags were defined as standing deadwood >10cm DBH (Spies et al., 1988). Girth was measured at a height of 1.3m from the base using measuring tape. Trigonometry was used to derive the heights of the snags using a clinometer and measuring tape. All snags within the 30x30m plot were measured. Snag decay class was assessed using a three-class system, identical to that described for DDW.

Soil pH and moisture measurements were taken in all of the 41 plots. Measurements were taken in each of the four corners using a Tecpel-707 (TECPEL Co. Ltd.), recording soil pH and soil moisture content. The sensitivity of the equipment could range from 3-8 pH, with an accuracy of ± 0.2 pH, and soil moisture content could range between 10-80%. It should be noted that pH and moisture readings were taken on different days for each of the plots, and thus may not be representative.

Hemispherical photography was used in order to estimate Leaf Area Index (LAI) and Canopy Closure (CC). Neither LAI nor canopy closure can be estimated accurately from a single image; Weiss et al. (2004) suggest a minimum of 8 images is required. In this study a total of 16 images were taken for each field plot at a height of 1m from the ground using a horizontally levelled digital camera. A uniform 4x4 cell grid was laid out in each field plot (see Figure 3.8) where a photograph was taken in each cell as close to the centre as possible,

allowing for obstructions. All photographs were taken between 10am and 2pm under overcast conditions to ensure a homogenous illumination of the overstorey canopy and a correct contrast between the canopy and the sky, as in Riano et al (2004).

Hemispherical photography was acquired using a Nikon Coolpix 5400 digital camera and Nikon FC-E9 fisheye converter (combined focal length equivalent to 7.2mm, and the combined F number is f/2.4). The field of view of this lens is approximately 183°. The Nikon Coolpix 5400 employed a high density 1/1.8" colour CCD sensor with a maximum pixel density of 13 MP/cm². Photographs were taken at a resolution of 2592 x 1944 pixels, illustrated in Figure 3.9. Each of the photos was oriented north to south in the field, according to the markers present on the self-levelling mount (SLM5, Delta-T Devices Ltd.), so north is always to the top-left of the image.

3.2.3 Airborne remote sensing

This section contains a summary of the remote sensing datasets acquired for the New Forest study site. Multi and hyperspectral imagery is outlined first, followed by discrete return and full-waveform LiDAR systems. Remote sensing datasets for the New Forest have been provided by the Airborne Research and Survey Facility (ARSF), a department of the Natural Environment Research Council (NERC). Bespoke data was acquired using the Dornier 228-101 research aircraft at an altitude of 5700ft (1737m), and speed of 135 knots (250 kph) for both April and July 2010. Thus, leaf-off and leaf-on datasets were acquired. A full listing of the remote sensing datasets is given in Table 3.4. Figure 3.10 illustrates the overlapping extents of the 2010 April and July remote sensing acquisitions.

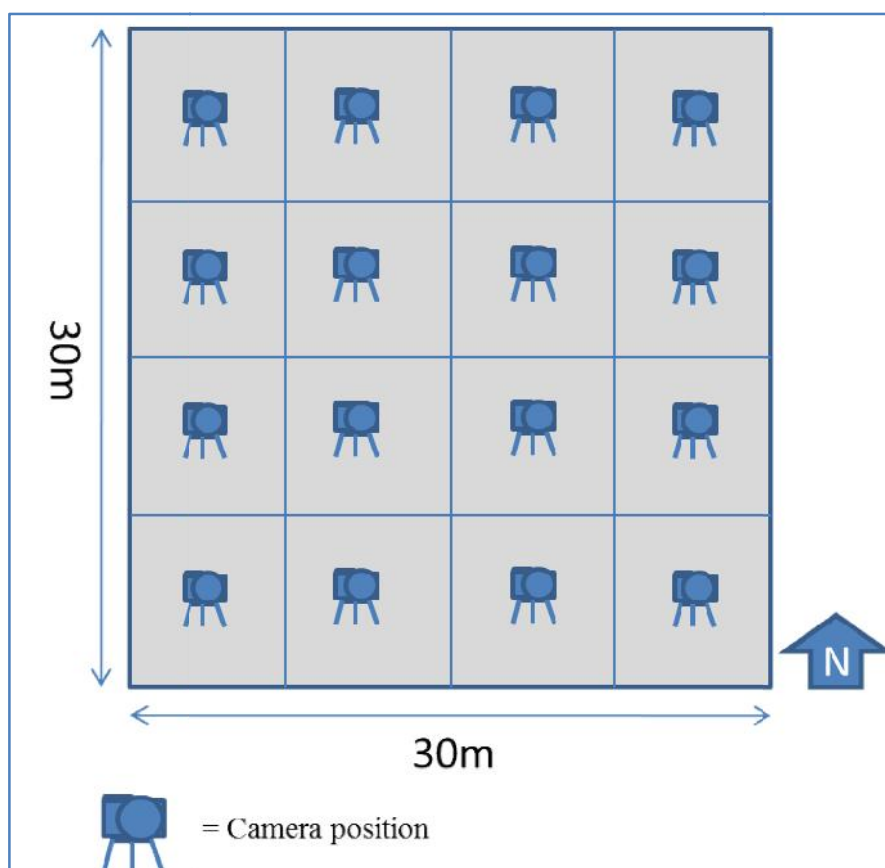


Figure 3.8 – A hemispherical photograph was taken as close to the centre of each of the 16 grid cells for each field plot site.



Figure 3.9 – Hemispherical photograph taken underneath semi-ancient deciduous woodland. The pointed marker (top-left) denotes north.

Table 3.4– A summary of all remote sensing datasets currently acquired for the New Forest.

Sensor name:	Description:	Resolution:
<u>July 11/07/2007</u>		
CASI-2	[Archived] 14 spectral bands (visible to near-infrared) set to MERIS specifications.	2x2m pixel size
<u>April 08/04/2010</u>		
Eagle	252 spectral bands (visible to near-infrared).	~1.1 x 1.1m pixel size Swath = 1013.05m
Hawk	233 spectral bands (shortwave infrared).	~2.1 x 2.1m pixel size Swath = 681.03m
ALS50-II LiDAR (DR&FW)	Discrete return (4 returns) and full waveform (2ns waveform sampling).	Discrete return at ~3.4 pulses returned per m ² . Waveform data at ~2.2 pulses per m ² .
<u>July 06/07/2010</u>		
Eagle	252 spectral bands (visible to near-infrared).	~1.1 x 1.1m pixel size Swath = 1013.05m
Hawk	233 spectral bands (shortwave infrared).	~2.2x 2.2m pixel size Swath = 681.03m
ALS50-II LiDAR (DR&FW)	Discrete return (4 returns) and full waveform (1ns waveform sampling).	Discrete return at ~3.7 pulses returned per m ² . Waveform data at ~2.4 pulses per m ² .

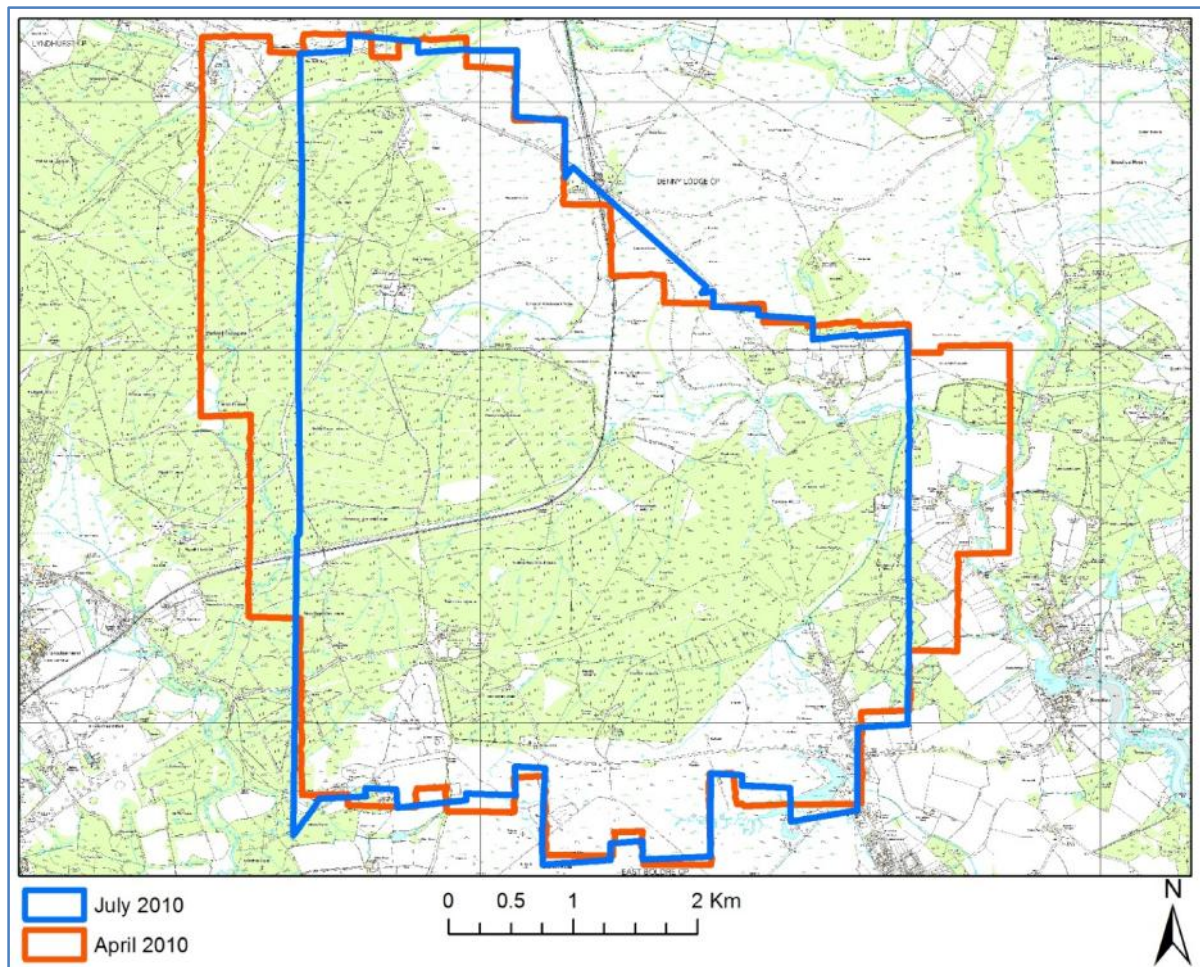


Figure 3.10 – The overlapping extents of the 2010 April and July airborne remote sensing acquisitions. Base Map layer is © Crown Copyright/database right 2010. An Ordnance Survey/EDINA supplied service.

3.2.3.1 Archived airborne multispectral data

The CASI-2 data acquired in 2007 were used to guide fieldwork only through a random stratification of the NDVI product. The CASI data were not used in any further analysis due to the relatively low spatial and spectral resolution and data acquisition being three years prior to this research project.

3.2.3.2 Airborne hyperspectral data

The ARSF acquired Specim AISA Eagle and Hawk hyperspectral datasets for leaf-off and leaf-on conditions in 2010 (April 8th and July 6th). The Eagle is a 12-bit, push-broom, hyperspectral sensor with a 1000 pixel swath width, covering the visible and near infra-red spectrum (400 - 970nm). For this research the nominal spatial resolution for the Eagle sensor was approximately 1.1x1.1m for both acquisition dates. The Eagle data contains 252 contiguous spectral bands at approximately 2.2nm bandwidths. The Hawk is a 14-bit sensor able to capture shortwave infrared wavelengths (970 - 2450nm). The Hawk has 320 spatial pixels, 244 spectral bands and a maximum spectral resolution of 8nm. The Hawk imagery was produced with a nominal spatial resolution of 2.1x2.1m for leaf-off and 2.2x2.2m for leaf-on acquisitions. The Hawk data contained 233 contiguous spectral bands, at approximately 6.2nm bandwidths. The dataset was distributed at level 1b, which has had radiometric calibration algorithms applied, to produce radiance or irradiance, and to which location and navigational information has been appended. These levels are named according to the National Aeronautics and Space Administration (NASA) standard product definitions.

The April and July Hawk data was not used in the course of this analysis due to errors in the data which could not be resolved due to sensor faults, in addition to the relatively coarse spatial resolution. Therefore only the leaf-on and leaf-off Eagle hyperspectral data were employed within this thesis.

3.2.3.3 Airborne LiDAR data (discrete and full-waveform)

The LiDAR data for this research were acquired in conjunction with the hyperspectral imagery by the ARSF. The LiDAR instrument used was the Leica ALS50-II airborne laser scanner. The system uses a scanning laser to measure the distance between the aircraft and the ground surface. The operating specification of the system includes a maximum pulse rate of 83 KHz, and up to 4 returns from each discrete laser pulse. The LiDAR was installed in the Dornier 228-101 research light-aircraft and flown over the area of interest. The position of

the aircraft whilst in the air was recorded using a differential GPS, with positional information being relayed from base stations in known positions. Orientations (roll, pitch, and bearing) of the aircraft were recorded through the Inertial Measuring Unit (IMU). The accuracy of geographical coordinates of the system is estimated at a maximum of ± 11 cm (including GPS errors) (Leica Geosystems, 2003).

In 2010, discrete return (DR) LiDAR data were captured at an altitude of approximately 5700ft (1737m), with a swath width of 612m across each flightline and a scan half-angle of 10° . Up to four returns could be extracted from each laser pulse. The DR LiDAR data provided from the ARSF were in LAS (version 1.2) file format, of X, Y and Z coordinates, intensity, and return number, for the first, intermediate, and last significant return per laser pulse. The X and Y location was supplied in British National Grid (BNG) coordinates, while elevation was supplied in metres above the Ordnance Survey of Great Britain 1936 Datum. In addition, a basic point classification of these points was provided by the ARSF, who used the Terrascan software (<http://www.terrasolid.fi>), to identify potential erroneous or noise points.

The Leica ALS50-II included an upgrade to record both discrete return and full-waveform (FW) LiDAR data for a small footprint simultaneously. The FW data was provided in LAS (version 1.3) file format, which contained GPS, IMU, and laser pulse return waveform data. For the April 2010 acquisition the ALS50-II scanner provided a measurement point density of 3.4 pulses per m^2 for discrete return measurements and approximately 2.2 pulses per m^2 for full-waveform measurements. For July 2010 data the measurement point density was 3.7 pulses per m^2 for discrete return measurements and approximately 2.4 pulses per m^2 for full-waveform measurements. The full-waveform measurements were of a lower pulse density due to the setup of the data capture and the additional processing time required to store the digitised waveform response. The full-waveform system allowed the full return waveform to be digitised and recorded with a digitisation sampling period of 1ns or 2ns. This sampling period defines the time intervals of sampling. The waveform is not completely recorded, but sampled over a predefined number of nano-seconds, 64, 128 or 256 (ARSF, 2011a). These samples are recorded from the time of the first return, with a small buffer before that (default is 5m) which in theory allows the lead-in to the first pulse. Both leaf-on and leaf-off full-waveform acquisition used 8-bit (or 256 nano-second) samples.

The 2010 leaf-off (April) waveform dataset was acquired with a 2ns sampling interval, whilst the 2010 leaf-on (July) dataset was acquired with a 1ns sampling interval. The sampling period imposes a number of peculiarities to the system, the first of which being the range resolution. Waveform digitisation sampling can be set to either 1 or 2ns (1ns is equivalent to 15cm distance travelled and 2ns equivalent to 30cm). The second peculiarity is that of maximum vertical range covered, where the digitiser starts recording from the first returned peak for the required sample number, e.g. 256, so this will define the maximum possible elevation of the last recording. The device was set to 256 samples at 2ns intervals for the leaf-off dataset, thus the maximum possible recorded height range per pulse is 76.8m from the first significant pulse (at nadir). For the leaf-on dataset the device was set to 256 samples at 1ns intervals, thus the maximum possible recorded height range per pulse is 38.4m.

The next chapter outlines the data processing and analysis methods applied to both the field data and the remote sensing datasets.

Chapter 4 –Methods

The following chapter concerns the processing steps involved in converting the raw datasets identified in Chapter 3 into data which can be analysed to achieve objectives 3-6 of this research project. To retrieve and assess metrics required to map woodland condition at the individual tree-level, plot- or stand-level, processing steps were developed for:

1. processing of field data to estimate critical plot-level variables and calculation of condition indices;
2. processing of hyperspectral datasets to generate usable individual tree and area-based metrics for analysis;
3. processing of DR LiDAR datasets to generate usable individual tree and area-based metrics for analysis;
4. processing of FW LiDAR datasets to generate usable individual tree and area-based metrics for analysis;
5. use of statistical methods to estimate comparable field-level attributes from the remote sensing data which would relate to condition assessment;
6. validation of the estimates for field plot-level data;
7. mapping the estimated field plot-level metrics across the study area and the calculation of condition indices from this data.

Figure 4.1, 4.2 and 4.3 provides flowcharts to illustrate the overall approach of this method to processing and analysing the two field datasets and the three remote sensing datasets (in addition to their combination) for predicting forest metrics necessary for assessing forest condition.

- Figure 4.1 depicts the basic processing steps necessary for fieldwork data analysis.
- Figure 4.2 illustrates the processing and analysis steps necessary for all remote sensing datasets.
- Figure 4.3 shows the necessary calculation steps for applying the field plot-level metric predictions across the whole study site in addition to the calculation of condition indices.

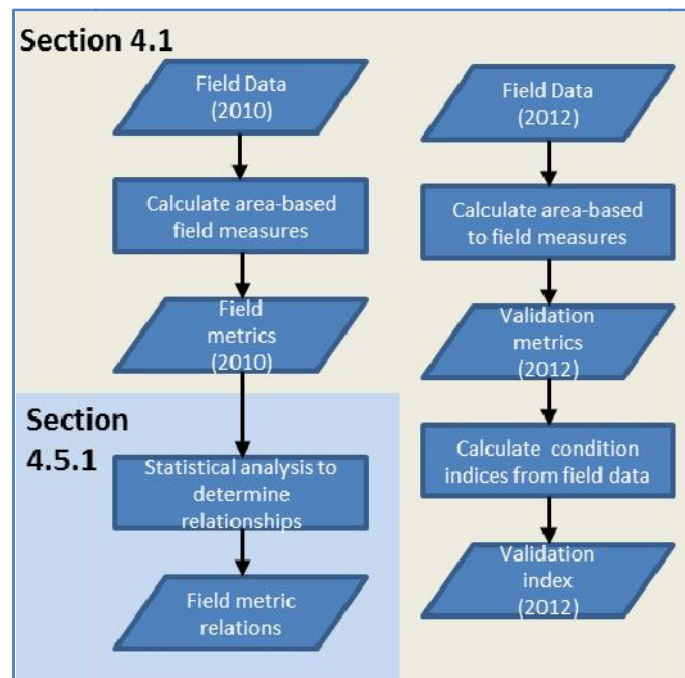


Figure 4.1 – A flowchart illustrating the processing and analysis steps for the field data

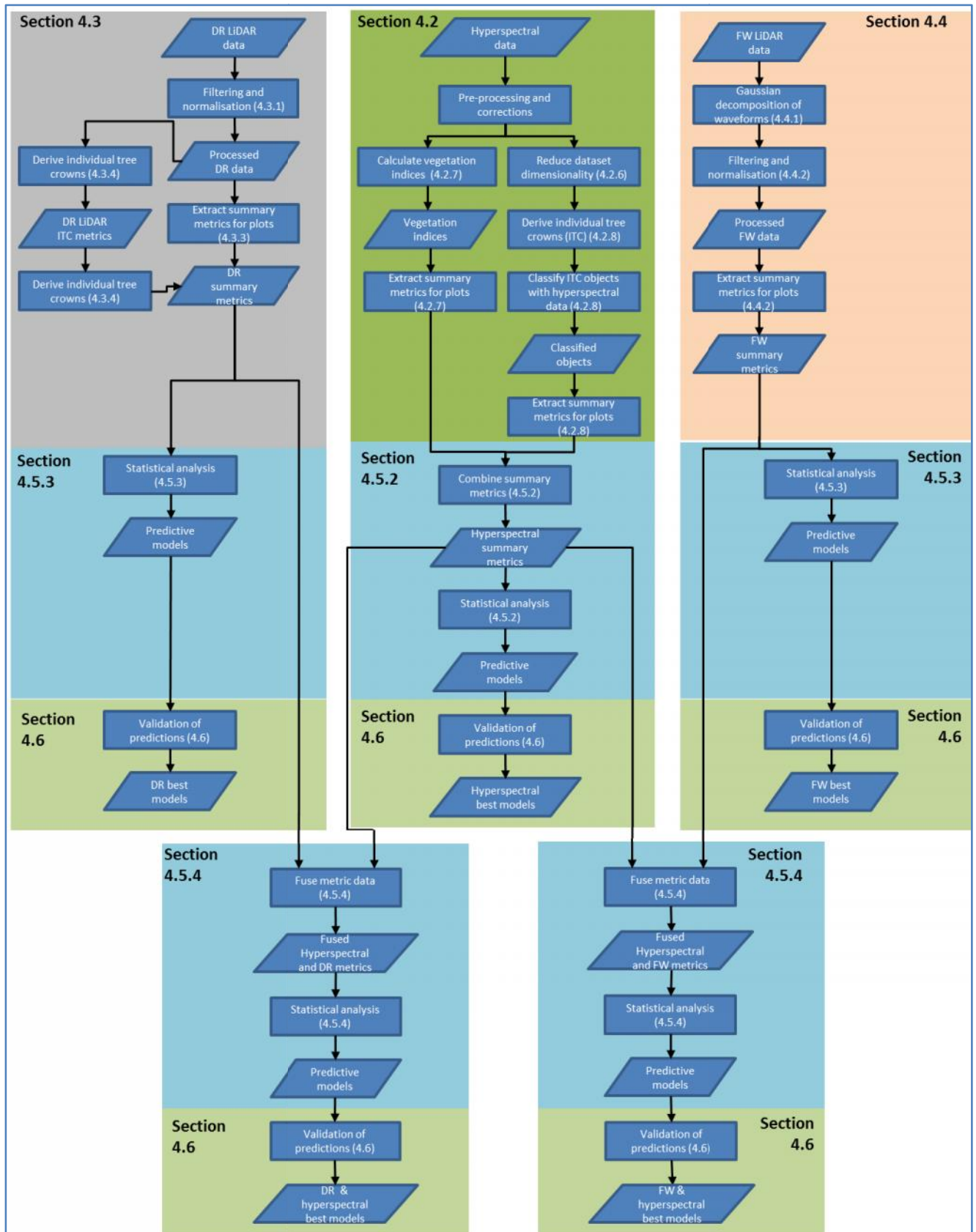


Figure 4.2 – A flowchart illustrating the analysis steps for remote sensing datasets.

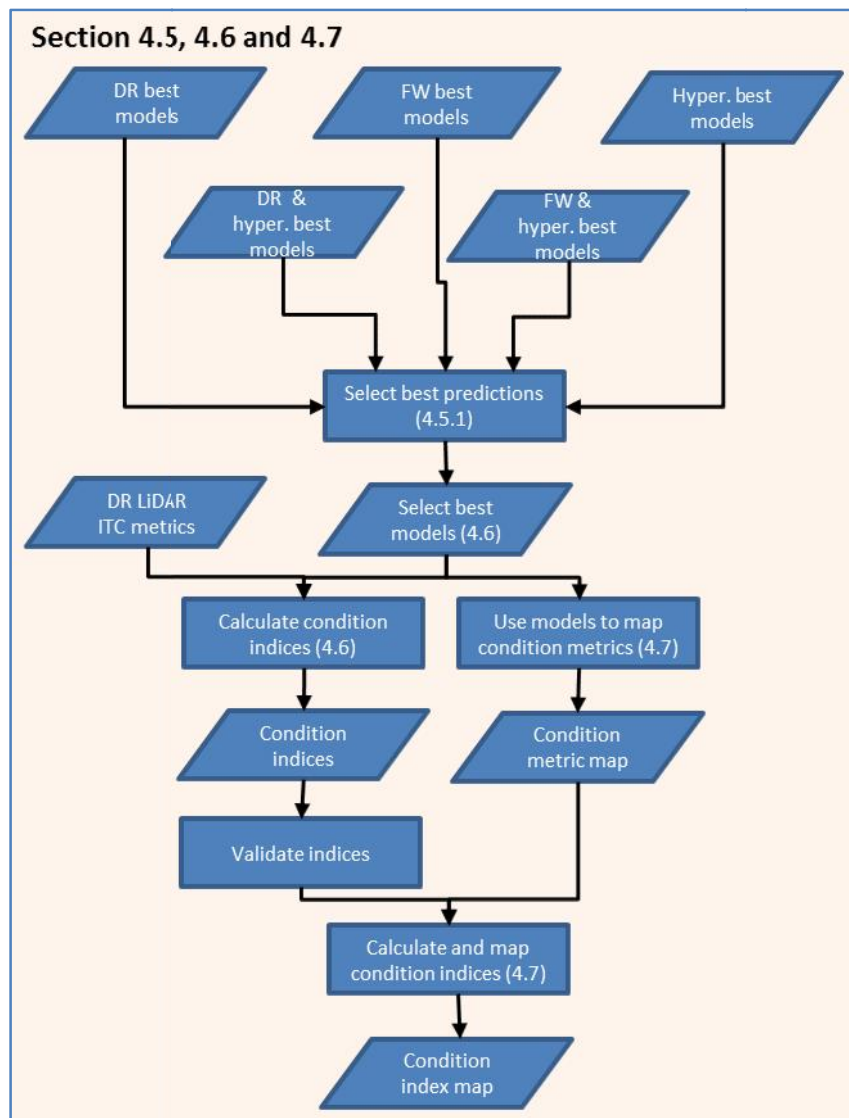


Figure 4.3 – Flowchart illustrating the selection of the best predictive models and the production of condition indices

4.1 Fieldwork analysis

What follows is a description of the calculation and processing steps applied to the data recorded in the field in order to provide the input required for the various condition indices. Choice of field metrics was selected through a broad search of the surrounding literature (see Chapter 2.2.4).

4.1.1 Calculations applied to the field datasets

A number of calculations needed to be applied to the fieldwork data in order to provide the required information for the various analysis steps outlined in section 3.2.2. Appendix A contains a full description of calculations applied to the fieldwork data; these variables are summarised in Table 4.1. The heights of trees and height of the first live branch (crown base) were calculated through the use of trigonometry from a combination of distance and angular measurements. Crown horizontal area was calculated through the use of the formula to derive the area of an ellipse. Diameter at Breast Height (DBH) and Basal Area (BA) were calculated using formulae for finding the diameter and area of a circle respectively from the circumference (i.e. tree girth). Deadwood volume was also calculated for snags and downed deadwood, using formulae for finding the volume of either cylinders or a truncated cone shape depending on the deadwood item (Cantarello and Newton, 2008). Each of the calculations mentioned were performed in Microsoft Excel 2007 (Microsoft Corp.).

4.1.2 Hemispherical photography processing

A total of 643 hemispherical photographs were taken in the field. Multiple images per site position were taken in order to experiment with exposure settings. Photographs were selected on the basis of the absence of blurs, distortions, saturation and the ability to discern tree features. A single photograph for each field plot division was chosen (see Figure 3.7). Each of the selected photographs was then input into Gap Light Analyser (GLA) software (version 2.0.4)(Frazer et al., 1999, 2001).

Figure 4.4 summarises the key steps for estimating canopy openness and Leaf Area Index (LAI) using GLA software. Image registration was applied manually by the user, selecting north as indicated on each the photographs from fibre optic markers, and defining the image area for analysis onto each photograph (see Figure 4.5). Figure 4.6 illustrates a sample photograph with the manually set threshold applied, in order to delineate tree canopy from

sky pixels. The intensities below the threshold were considered to be foliage (black), whilst those above are classified as sky (white) (Frazer et al. 2001). This process was subjective.

Canopy openness (the opposite of canopy closure) is a sine-weighted measure which represents the relative amount of unobstructed (open) sky visible from an unobstructed point in the understorey (Frazer et al. 1999). LAI is a dimensionless variable, the ratio of leaf area per unit of ground surface area. This indirect method of LAI estimation uses a modified Beer's Law light extinction model, by mathematically analysing the light intercepting effect of leaves with different angular distributions using a simplified assumption of randomly distributed foliage elements. LAI is discussed in greater detail in Zheng and Moskal (2009).

Table 4.1 – Summary of calculations applied to field recorded data (more detailed descriptions in appendixA).

Calculation:	Description:
Number of tree stems per hectare	The field plot size corresponds to 9% of 1 ha. The total number of tree stems was extrapolated accordingly
Number of saplings and seedlings per hectare	The number of native and total saplings and seedlings were totalled and extrapolated to their occurrence at 1 ha.
Diameter at breast height (DBH) (m)	Estimated from girth measurement for individual trees. A mean and standard deviation were also calculated for each plot. Plot mean and standard deviation were then calculated.
Percentage of big trees per plot	Those trees with DBH between 0.4-0.8 and DBH>0.8 were counted relative to the total number of trees within the plot as a percentage.
Basal area (BA) (m ²)	BA is essentially the horizontal area taken up by the tree trunk. This was calculated for each tree and summed to the plot-level. An estimate of BA per hectare was produced.
Tree heights (m)	Derived through trigonometric calculations using tape and clinometer measurements for total tree height.
Lorey's mean height (m)	Lorey's mean height weights the contribution of trees to the plot height by their basal area.
Height to the first live branch (m)	Derived through trigonometric calculations using tape and clinometer measurements for height from base of the stem to first live branch. Plot mean and standard deviation were then calculated.
Crown horizontal area (m ²)	To estimate the horizontal coverage of the crown the North-South and East-West extents were input into the formula to estimate the area of an ellipse. Plot mean, standard deviation and total were then calculated.
Volume of downed deadwood (DDW) (m ³)	Includes fallen trees and fallen branches. Volume was estimated from length and diameter measurements.
Volume of snags (m ³)	Standing dead trees only. Volume was estimated through height (assessed through the same method as for tree height) and girth measurements.
Number of ground flora species per hectare	Concerning vascular plants. The total number was counted. The subplot size corresponds with 1% of 1 ha. The total number was extrapolated.
Sum of DBH differences between neighbouring trees	Input to the Tree Diameter Differentiation index (Vorčák et al., 2006). Sum of DBH differences between neighbouring trees.
Index of vertical structure	Part of the Stand Diversity Index (Jaehne and Dohrenbusch, 1997, in Newmann and Starlinger, 2001, Vorčák et al., 2006). The ratio of the three smallest DBHs against the three largest DBH values within the plot.

The LAI index value can range from 0 for bare ground to 6 for a dense forest. Two types of LAI estimation were made, those calculated from LAI ring 4 and LAI ring 5. LAI ring 4 is the effective LAI calculated over zenith angles 0 to 60° from the centre of the hemispherical photograph (i.e. the central 30% area of the image). LAI ring 5 is the effective LAI integrated over zenith angles of 0 to 75°. Estimates of canopy openness and the two LAI approximations were recorded for each hemispherical photograph, and the values averaged for each field plot using Microsoft Excel 2007.

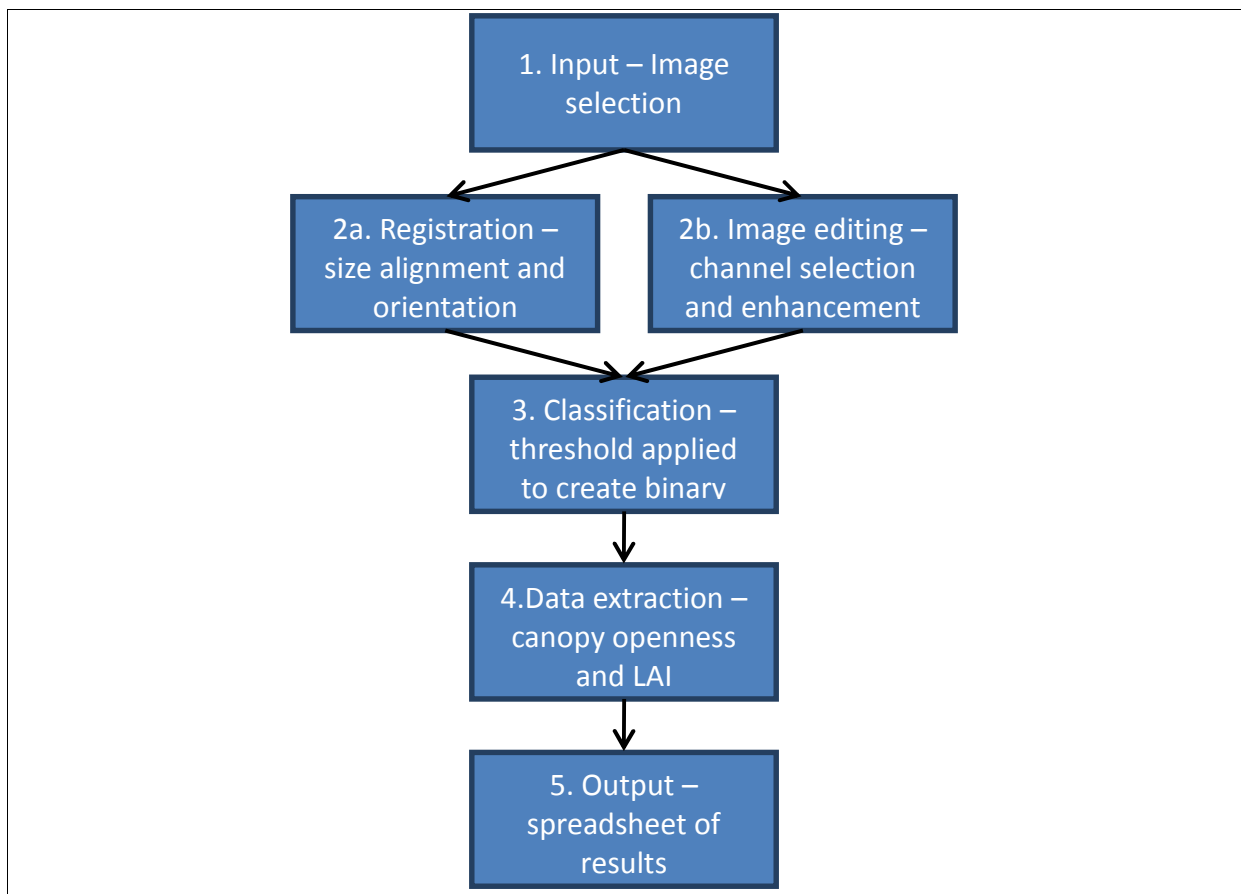


Figure 4.4 – The 5 key steps to analysing hemispherical photographs for forest canopies.

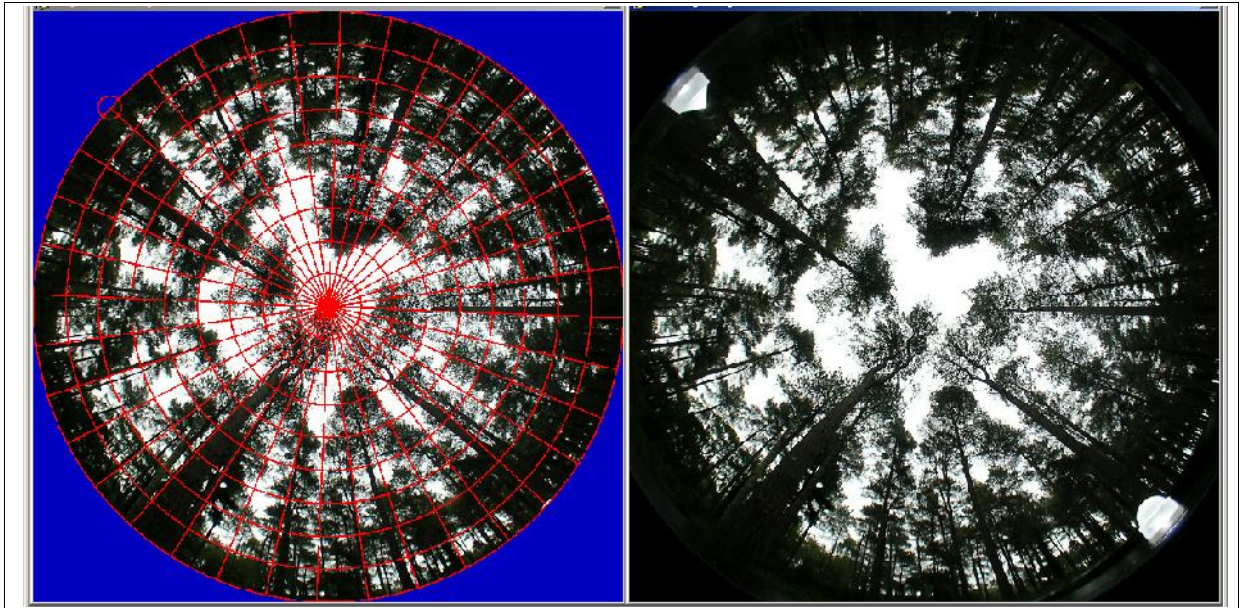


Figure 4.5 – (Left) Registered image, note the red circle in the top left corner identifies geographic north (0°). (Right) the original photograph with north and south marked at time of exposure using fibre optics mounted at the edge of the fish-eye lens.

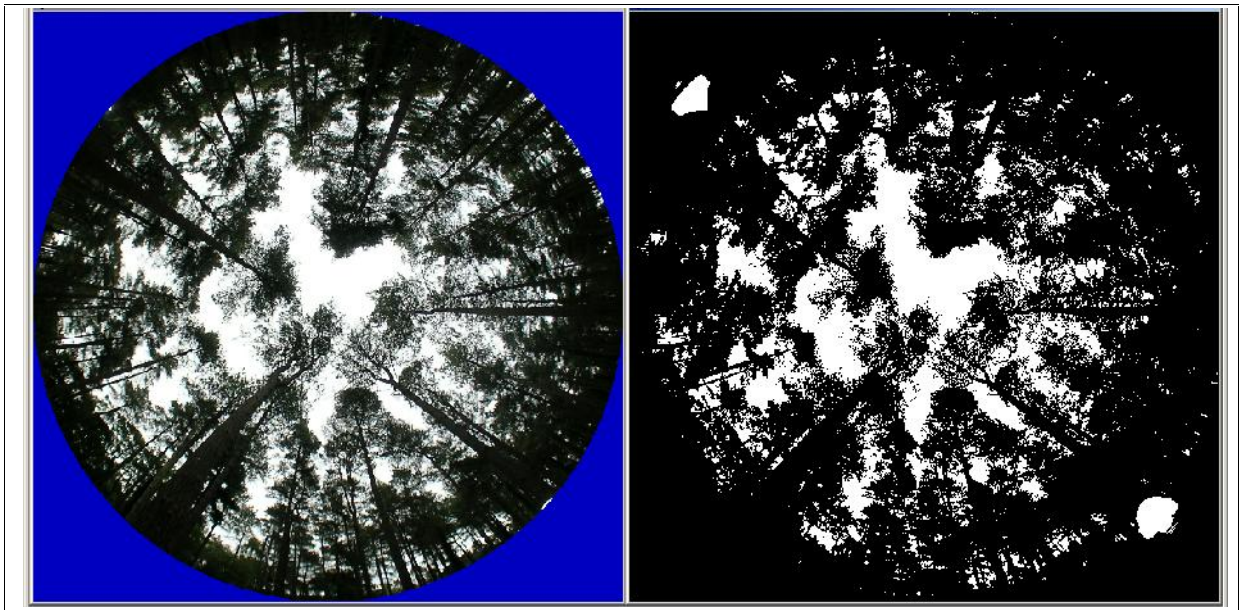


Figure 4.6 – (Left) The original photograph, (right) the binary threshold image, where foliage is considered black and sky as white.

4.1.3 Calculating compositional and structural indices

With the field data available, nine measures for both species and structural diversity were calculated for each plot. Thus, all the indices refer to within plot structure and composition. There are a considerable number of other indices for describing spatial structure (see section 2.2.2). Further examples are available in Pommerening (2002, 2006) and Staudhammer and Lemay (2001).

Only a subset of indices was explored within the course of this project, the choice of which was determined by their number of inputs and computational complexity. Some indices used focus on a single facet of the woodland environment, such as the spacing of tree stems, whereas others utilise multiple facets such as species composition, tree size, size variability, the presence of deadwood and the regeneration of species. In addition some indices require the input of area-based or individual tree-based metrics. The aim was to test for which of the techniques provide a better woodland condition indicator. What follows is a description of the indices chosen for this research project.

4.1.3.1 Species composition indices

The species compositional indices of the Shannon-Wiener index (SH) (Shannon, 1948) and the Simpson index (SI) (Simpson, 1949) were utilised in this study. These indices are a mathematical measure of species diversity within a community. Both SH and SI were calculated with the proportion of stems of that species within the plot, and the proportional basal area of the species within the plot. In addition a measure of Evenness (E) was computed, which is a measure of how similar the abundances of different species are. The SH has been criticised however because it is sensitive to sampling size and over-emphasises rare species (Magurran, 2004). An alternative is the SI. A combination of indices was chosen to take advantage of the strengths of each and develop a better understanding of species composition.

The SH diversity index for all tree species, and native seedling species was calculated as:

$$SH = - \sum_{i=1}^n p_i \log_e p_i \quad [4.1]$$

Where p_i = the proportion of individuals (trees count/seedlings count/total basal area) in the i th species, and n is the number of species. The Shannon index was calculated for stem number for tree and seedlings species, in addition to basal area for tree species. Typically the

value of the index ranges from 1.5 (low species richness and evenness) to 3.5 (high species evenness and richness).

The Evenness index (E) removes the influence of species number by standardisation, and is performed upon the Shannon-Wiener Index (as above), it is calculated as:

$$E = SH / \log_e n \quad [4.2]$$

E is a measure of how similar the abundances of different species are. When there are similar proportions of all subspecies then $E = 1$, but when the abundances are very dissimilar (some rare and some common species) then the value increases.

The Simpson index was calculated for tree species in each plot as:

$$SI = 1 - \left(\sum_{i=1}^n (1 - p_i^2) \right) \quad [4.3]$$

Note this is the ‘Complementary’ form of the index. Index values can range between 0 and 1, where diversity increases, the index value increases. It should be noted that using the Simpson index with low numbers, misleading results can be obtained. A low Simpson index value equates to high diversity, whereas a high value correlates to a low diversity.

4.1.3.2 Structural indices

The indices for spatial distribution or horizontal structure compare a hypothetical distribution with the real situation. The first of these is the Clark-Evans aggregation index (R) (Clark and Evans, 1954). This index was developed for the purposes of botanic and phytocoenologic studies. In forestry research it has been used very rarely (Vorčák et al., 2006). The aggregation index describes the horizontal distribution of trees using the relation of the mean distance between the reference tree and its nearest neighbour and the expected distance between them for a random distribution in the stand. This is calculated as follows:

$$R = \frac{\frac{1}{N} \sum_{i=1}^N r_i}{0.5 \times \sqrt{\frac{Pl}{N}}} \quad [4.4]$$

Where r_i is the distance of tree i to the nearest neighbour, N is the number of trees on the sample plot, and Pl is the area of the sample plot (m^2). The aggregation index can theoretically range from 0 at maximum tree clustering, to 2.15 at the regular hexagonal distribution of trees. An index value of 1 indicates that the trees are distributed randomly

within the stand. Thus, stands with the index value > 1 show a tendency towards regular distribution, while values of < 1 show a tendency towards clustering.

Vorčák et al. (2006) state however that it is not enough to know the index value alone. In nature most variables have a stochastic character and this index belongs to this type of variable. It is therefore important to know whether the difference between the calculated index and the index expected from a random distribution is significant. This fact can be tested with:

$$t = \frac{r_R - r_T}{\sigma_{rT}} = \frac{r_R - r_T}{\frac{1.26136}{\sqrt{N \times \frac{N}{Pl}}}} \quad [4.5]$$

Where r_R is the real distance between the tree and its nearest neighbour, r_T is the expected distance to the nearest neighbour, σ_T is the standard error of the mean for the expected distance to the nearest neighbour, N is the number of trees on the sample plot and Pl is the area of the sample plot (m^2). If the t-value is higher than 1.96 with 95% probability, the trees in the stand have a clustered or regular distribution according to the value of the R for the given stand.

Tree diameter differentiation has been identified as another potentially important parameter of structural diversity, for which an index (TM) was created by Földner (1995 cited in Vorčák et al., 2006). It can be calculated from various tree attributes (diameter, perimeter, basal area, height, or volume). Diameter is the most commonly used attribute. The differentiation can be quantified by the following formula:

$$TMi = \frac{1}{N} \sum_{j=1}^N (1 - d_{ij}) \quad [4.6]$$

Where N is the number of trees on the sample plot, and d_{ij} is the relation between thinner and thicker DBH in the analysed neighbour tree pair. The index values range from 0 to 1. The stands with small diameter differentiation have the index values near 0, whilst stands with high diameter differentiation reach index values closer to 1. Aguirre et al. (1998 cited in Vorčák et al., 2006) suggested a 5-level scale in order to simplify the comparison of the stands as follows: low differentiation (0.0-0.2); medium differentiation (0.2-0.4); obvious differentiation (0.4-0.6); strong differentiation (0.6-0.8); and very strong differentiation (0.8-1.0).

There are comparatively fewer indices for the measurement of vertical structure (Neumann and Starlinger, 2001). Neumann and Starlinger (2001) outline the vertical evenness index (VE), in order to characterise the vertical distribution of coverage within a stand. The trees within the plot were stratified into four layers (100-80%; 80-50%; 50-20%; 20-0% of the maximum height of the plot), their crown projection area was then calculated, and the Shannon formula applied to the resulting proportions. Furthermore the results were standardised by the four layers according to the Evenness formula. Thus the equation is:

$$VE = \frac{\sum_{i=1}^4 \pi_i \log \frac{\pi_i}{\log 4}}{\log 4} \quad [4.7]$$

Where π_i is the relative crown area of all trees in the i th height layer. Low values of VE characterise single storied stands while the theoretical maximum of 1 would result for forest canopies which contained a high degree of different vertical layers.

4.1.3.3 Combined indices

All of the indices documented so far concentrate upon single elements of diversity. The complexity index (HC) by Holdridge (1967) relies on combining traditional measures of stand description:

$$HC = \frac{(H \times BA \times N \times n)}{10000} \quad [4.8]$$

where multiplying the number of trees species (n) by stem number per ha (N), dominant height (H), and basal area per ha (BA), divided by the field plot area. The index is determined strongly by the number of species and measures growth performance but contains no information on spatial distribution nor accounts for within-stand variation. Typically, a target index value is assigned from an ‘ideal’ woodland and other sites are assessed against this value (McElhinny, 2005).

Jaehne and Dohrenbusch (1997, cited in Neumann and Starlinger, 2001, Vorčák et al., 2006) proposed the complex stand diversity index (CSDI), which was stated as integrating the main elements of diversity by combining measures for the variation of species composition (A), vertical structure (S), spatial distribution (V), and crown dimensions (K) into a single figure. This index can be presented as:

$$B = p \times (A + q \times S + V + K) \quad [4.9]$$

where p and q are factors of importance ($p=4$ and $q=3$). Table 4.2 describes each of the four variables of stand structural diversity. Jaehne and Dohrenbusch (1997, Vorčák et al., 2006)

also offered the general evaluation of the stand diversity according to the *CSDI* value, these are:

1. $B \geq 9.0$ = a very heterogeneous stand structure;
2. $8.0 \leq B < 8.9$ = heterogeneous stand structure;
3. $6.0 \leq B < 7.9$ = uneven stand structure;
4. $4.0 \leq B < 5.9$ = homogenous stand structure;
5. $B < 4.0$ = monotonous stands.

Table 4.2 – Components of the CSDI of stand structural diversity.

Component:	Description:
1. Index of tree species composition (<i>A</i>)	$A = 1 + \log \left(\frac{Z - Ma_{max} + Ma_{min}}{n} \right)$ <p>Where n is the number of species; Z is a control parameter (the authors suggest the value 1.5); Ma_{max} is the relative proportion of the most abundant tree species; and Ma_{min} is the relative proportion of the least abundant tree species.</p>
2. Index of vertical structure (<i>S</i>)	$S = 1 - \frac{\sum_{i=1}^N DBH_{Hmin}}{\sum_{i=1}^N DBH_{Hmax}}$ <p>Where N is the number of measured trees (3 thickest and 3 thinnest trees); DBH_{min} is the DBH of the thinnest trees (in cm); and DBH_{max} is the DBH of the thickest trees (in cm).</p>
3. Index of spatial distribution (<i>V</i>)	$V = \left(1 - \frac{\sum_{i=1}^n Ab_{min}}{\sum_{i=1}^n Ab_{max}} \right) \times f \times st$ <p>Where n is the number of measured distances (3 shortest and 3 longest distances between neighbour trees); Ab is the distance between trees (m); and f is the correction for stand density (in the pole stage and older stands this can be omitted);</p> $f = Y \frac{1}{\frac{\sum_{i=1}^n Ab_{min}}{n} + \frac{\sum_{i=1}^n Ab_{max}}{n}}$ <p>st is the factor considering coppice sprouts;</p> $st = N_{250} \times 0.1 + 1$ <p>N_{250} is the number of coppice sprouts per 250m²; and Y is the control parameter.</p>
4. Index of crown differentiation (<i>K</i>)	$K = \left\{ 1 - \log \left(\frac{\sum_{i=1}^n Ka_{min}}{n} \right) \right\} + \left(\frac{\sum_{i=1}^n Kd_{min}}{\sum_{i=1}^n Kd_{max}} \right)$ <p>Where n is the number of selected trees (2 trees with the smallest and 2 trees with the largest crown diameter); Ka_{min} is the smallest height to the crown base (m); Kd_{min} is the smallest crown diameter (m); and Kd_{max} is the largest crown diameter (m).</p>

4.1.3.4 Score-based indices

The scoring method was defined by Cantarello and Newton (2008) and requires the definition of target indicator values for the forest type in question. Their study was focused upon old growth deciduous woodland in the New Forest, UK, rather than a combination of plantation coniferous and deciduous woodlands. Thus, all the targets used in this approach were typically based upon ‘ideal’ deciduous forest conditions. After a review of various biodiversity indicators was undertaken for habitat monitoring under the Natura 2000

initiative, 17 indicator variables were selected for the investigation, see Table 4.3. Those indicators were reported as the most commonly featured in the literature, focusing primarily on indicators of forest structure and function which have been tested at the local scale. Field plot measures were assessed against the target values and assigned a 0 if lower and 1 if equal to or above the target. The sum of the scores for the indicators represents the index. This index was aimed mainly at assessing old-growth deciduous woodlands; predominantly the targets for deciduous trees were based upon published figures for Denny Wood, New Forest (Mountford et al., 1999)

Table 4.3 – List of ecological indicators analysed and their target values for the New Forest

Key Factor	Indicator	Target
Forest Stand Structure	1. Number of trees (ha^{-1})	222 ^{*3}
	2. Shannon-Wiener index for native trees	0.87 ^{*3}
	3. Basal area (m^2ha^{-1})	23 ^{*3}
	4. Mean DBH (cm)	32 ^{*3}
	5. Standard deviation of diameters (cm)	14 ^{*4}
	6. Percentage of big trees	7 ^{*3}
	7. Mean height (m)	17 ^{*1}
	8. Number of total saplings (ha^{-1})	91 ^{*3}
	9. Number of native saplings (ha^{-1})	91 ^{*3}
Deadwood	10. Volume of downed deadwood (m^3ha^{-1})	26 ^{*3}
	11. Downed deadwood decay class	0.5 ^{*2}
	12. Volume of snag (m^3ha^{-1})	16 ^{*3}
	13. Snag decay class	0.4 ^{*2}
Tree Regeneration	14. Number of seedlings (ha^{-1})	63,219 ^{*3}
	15. Number of native seedlings (ha^{-1})	63,219 ^{*3}
	16. Shannon-Wiener index for native seedlings	0.89 ^{*3}
	17. Number of ground vegetation species	33 ^{*3}

^{*1}Ferretti et al. (2006); ^{*2}Keddy and Drummond (1996). ^{*3}Mountford et al. (1999); ^{*4}Van den Meersschaut and Vanderkerkhove (1998);

4.2 Hyperspectral data processing

This section describes the pre-processing steps required in order to allow any form of analysis with the hyperspectral datasets. Only the Eagle leaf-off (April) and leaf-on (July) datasets acquired in 2010 were used for this part of the processing.

4.2.1 Hyperspectral data pre-processing

The hyperspectral datasets were supplied in level 1a format by the ARSF according to the standard NASA definition (ARSF, 2011b). These data cannot readily be used for analysis. The following pre-processing steps were required to generate level 3b data, as defined by the ARSF (2011b), for analysis purposes. For example, a number of pre-processing steps were required in order to correct the hyperspectral imagery for illumination, geometric and atmospheric effects. Each of the flightlines was a single file and required merging with its

neighbours in order to create a continuous coverage. In addition, steps to eliminate noise and extraneous data were implemented. The following sections describe in detail each of these necessary pre-processing steps in order.

4.2.1.1 Cross-track correction

Upon visual examination of each of the Eagle flightlines a limb-brightening effect was observed. This effect will have been caused by both path length effects and differential scattering angles of atmospheric aerosols (Hill et al. 2009a). The hyperspectral data were acquired near to mid-day in order to maximise solar irradiance and signal strength, with the majority flown in a north-south orientation. This orientation resulted in minimised limb-brightening, by reducing aerosol optical depth which was symmetrical about the nadir-line of the flightline. For the few flight strips flown perpendicular to the solar azimuth, limb-brightening was at its greatest and asymmetric as a result of the larger scattering angle between the solar zenith/azimuth angle and the across scan angle ($\pm 45^\circ$).

Radiometric corrections were applied using ENVI 4.7 (ITT Visual Information Solutions) to minimise the limb-brightening effect and reduce the general upward trend in spectral response towards the edge of each scan line. Along-track mean values were calculated and were used to show the mean variation in the cross-track direction. This captured the overall limb-brightening for each flightline. Third order polynomial functions were used to fit the means and remove the illumination variance across the image (ENVI-Online-Help 2005). This minimised the limb-brightening effect while retaining scene variability. This step was required before geo-corrections could be applied.

4.2.1.2 Geo-corrections

The level 1a hyperspectral flightline data contained no location information to match pixels to a known coordinate system. Thus, hyperspectral imagery was geometrically corrected utilising an Azimuth Systems program AZGCORR (<http://arsf.nerc.ac.uk/data/azimuth.asp>) which was used to (i) apply aircraft navigation data to each scan line and (ii) project this data upon a geoid-based projection to determine the exact intersection of each pixel's view angle with a high spatial resolution Digital Surface Model (DSM). The DSM used here had a 1m spatial resolution and was generated from the concurrently captured LiDAR data (see section 4.3). The product of this was a geo-correction aligned to the British National Grid (BNG).

The Eagle imagery was re-sampled to 1.1m pixel spatial resolution utilising a bilinear interpolation algorithm.

Each of the geo-corrected flightlines for each of the dates was checked visually for geometric accuracy by overlaying vectors of road edges, building outlines, and water features from the Ordnance Survey Meridian 2 dataset, as demonstrated by NCAVEO (2005). The ARSF also checked the geospatial integrity of the data in this manner prior to delivery for this project (ARSF 2011b), and stated that the hyperspectral data corresponds spatially to within 0-2m of the Ordnance Survey vector overlay maps. This process was repeated here and the registration between the hyperspectral data and OS vectors was confirmed as being within 2m.

4.2.1.3 Atmospheric correction

The solar radiation reflected from the Earth's surface must pass through part of the atmosphere before it is collected by the airborne instrument. Thus the recorded data contains information about the atmosphere as well as the Earth's surface. For a quantitative analysis of surface reflectance to be performed, the removal of the influence of the atmosphere is a critical pre-processing step. The properties of water vapour and aerosols within the atmosphere are rarely known; but there are techniques to infer them from their imprint on hyperspectral radiance data (ENVI-Online-Help 2005). These properties can then be used within highly accurate models of atmospheric radiation transfer to produce estimates of true surface reflectance. Atmospheric corrections can be applied on a pixel-by-pixel basis.

It should be noted that the atmospheric correction step is necessary as all vegetation indices require high-quality reflectance measurements from either multispectral or hyperspectral sensors. Measurements in radiance units that have not been atmospherically corrected are unsuitable, and typically provide poor results (San and Suzen, 2010).

As close range calibration data were not available for all datasets, a repeatable solution for both 2010 Eagle datasets was required. The atmospheric correction model in ENVI 4.7 was used for this task, i.e. FLAASH (Fast Line-of-sight Atmospheric Analysis of Spectral Hyper-cubes), which is a first-principles atmospheric modelling tool for retrieving spectral reflectance from hyperspectral radiance images. FLAASH incorporates the MODTRAN 4 radiation transfer code; for more details on this code see Matthews et al. (2000). Atmospheric

corrections of this type can be applied on a pixel-by-pixel basis because each pixel in a hyperspectral image contains an independent measurement of atmospheric water vapour absorption bands.

FLAASH operates from the standard spectral radiance equation, as described in Matthews et al. (2000). The calculations are based on the viewing and solar angles and the mean surface elevation measurement, assuming a certain type of atmosphere, aerosol level and visibility. Corrections are made for water vapour column amount by analysing selected wavelength channels to retrieve estimated amounts for each pixel.

Close range spectral radiometry was only available for the April 2010 data capture, and as such empirical line correction could not be applied to both sets of imagery. The ARSF Field Spectroscopy Facility (FSF) in Edinburgh (<http://fsf.nerc.ac.uk/>) loaned a number of pieces of equipment for this task. These included four reflectance tarpaulins to provide a uniform spectral target for detection by the airborne imagery and for the following device. The Spectra Vista Corp (SVC) HR-1024 Field Spectroradiometer records the spectra of surface material within the visible to short-wave infrared wavelength range (350-2500nm). This was utilised to sample the reflectance tarpaulins and a number of other vegetated and ground cover types, most notably gravel, leaf-litter, grassland, heathland and various tree and heathland vegetation types. In addition, the Microtops II Sunphotometer, a hand-held instrument for measuring the atmospheric column, was also available. This device is used for the determination of aerosol optical thickness (AOT), utilising filters at 936nm and 1020nm for the measurement of total water vapour in the atmospheric column. A reading on this device was taken every five minutes while the ARSF data acquisition was under-way.

The data recovered from the fieldwork in April 2010 was used for assessing the accuracy of the FLAASH corrections. Calculated radiance values were checked against values recorded in the field.

4.2.1.4 Between flightline radiometric difference correction

In the 2010 Eagle data some minor variation in pixel reflectance values was evident when comparing adjacent flightlines. These differences were caused by changing levels of solar illumination of the study area during the time of the flight, primarily caused by changing

positions of cloud or by the automated changing of settings with the sensors on the aircraft. In addition there may have been residual errors in the cross-track illumination corrections.

A histogram matching approach was implemented in order to account for these differences. This function within ENVI 4.7 makes the brightness distribution of two flightlines as close as possible through a method of stretching the pixel values of one image to approximate another for each spectral band (ENVI-online-help, 2005). Errors unable to be mitigated (i.e. cloud shadow areas) were manually digitised and masked out of the flightline prior to mosaicking.

4.2.1.5 Image mosaicking

Image mosaics were produced from the hyperspectral datasets in order to produce a continuous high spatial-resolution surface. The Eagle images were mosaiced according to their geo-reference information within ENVI 4.7. The geometric and radiometric accuracy of the airborne remotely sensed data are better towards the central nadir viewpoint of the individual flightlines (Hill et al. 2009a). Thus, subsetting and the use of feathering options were used within ENVI to preserve these near-nadir areas where possible. Preference was given to the north-south flightline images to ensure similar illumination characteristics. Those areas where clouds or illumination anomalies had been masked out and removed in one flightline were filled with data from an overlapping flightline.

4.2.1.6 Quantification of horizontal accuracy

Consideration must be made with regards to spatial accuracy and co-registration with other datasets. A total station was utilised to provide accurate positional information for the training and validation field plot locations. Inaccuracy with spatial locations would introduce additional uncertainties in being able to relate field measurements with remote sensing metrics.

As the spatial accuracy of the LiDAR data was considered sufficient, as defined in section 4.3, the geometric accuracy and co-registration of data layers was tested by calculating the displacement error of features in the hyperspectral dataset compared with those in the rasterised LiDAR digital elevation model (DEM).

Error was estimated by using the georeferencing tool within Arcmap (version 10) by assigning 20 ground control points (GCP) to the hyperspectral image and DR LiDAR derived

raster DEM. A first order polynomial transform was applied to these GCPs to compute the route mean square error (RMSE) between the two datasets. For the leaf-on Eagle hyperspectral dataset assessed against leaf-on DR LiDAR the total RMSE was calculated as 1.38m. The leaf-off Eagle dataset assessed against the leaf-off DR LiDAR had a total RMSE calculated as 1.42m

4.2.1.7 Image dimensionality reduction

The hyperspectral dataset covered a large spatial extent in addition to a large, and finely defined, spectral extent. The following defines the pre-processing steps necessary to combine the datasets prior to analysis.

The Eagle imagery has a high number of spectral bands, often referred to as a high-dimensionality dataset. The benefits of such a dataset can be the better discrimination among similar spectral signatures than the more common multispectral sensors with limited spectral resolution (Huang and He, 2005). However the vast amount of data presents a challenge for information processing in terms of heavy computational burden and problems in storing such datasets, known as the Hughes phenomenon (Hughes, 1968). The image dataset may contain redundancies. Here these are bands whose reflectance correlates significantly with another band. The image bands may also contain noise. The definition of noise is: signal that does not correlate at all between bands.

The purpose of the dimensionality reduction was to reduce the complexity of tree species classification and classification process execution time. Within ENVI 4.7 there are two potential methods of dimensional reduction, each of which produce uncorrelated output bands, segregate noise components, and reduce the dimensionality of datasets. These are the Principle Component Analysis (PCA), and the Minimum Noise Fraction (MNF) transforms. Table 4.4 summarises the two procedures. A full description of theory behind these approaches is available in Chang (2007).

Table 4.4 – Methods of reducing image dimensionality in ENVI (version 4.7)

Name:	Description:
Principle Component analysis (PCA)	<ul style="list-style-type: none"> PCA is used to produce uncorrelated output bands. This is accomplished by finding a new set of orthogonal axes that have their origin at the data mean and that are rotated so the data variance is maximized. The resultant PCA bands are linear combinations of the original spectral bands and are uncorrelated. The same number of output PCA bands can be calculated as input spectral bands. The first PCA band contains the largest percentage of data variance and the second PC band contains the second largest data variance, and so on. The last PCA bands appear noisy because they contain little variance, much of which is due to noise in the original spectral data (Richards, 2012).
Minimum Noise Fraction (MNF)	<ul style="list-style-type: none"> A minimum noise fraction (MNF) transformation is used to reduce the dimensionality of the hyperspectral data by segregating the noise in the data. The MNF transform is a linear transformation which is essentially two cascaded Principal Components Analysis (PCA) transformations Green et al. (1988). The first transformation de-correlates and rescales the noise in the data. This results in transformed data in which the noise has unit variance and no band to band correlations. The second transformation is a standard PCA of the noise-reduced data. [ENVI-online-help, 2005]

The PCA and MNF transformations were applied to both leaf-on and leaf-off Eagle hyperspectral data. Each approach provided up to ten image-bands per dataset. The pixel values contained within each of the bands could not be related to any specific spectral measurements after the transform. Each of the resultant image-bands was inspected for the presence of noise and artefacts. If the image-band was considered too noisy it was removed.

PCA and MNF transforms were applied to the leaf-on and leaf-off datasets, which produced summary statistics for each image band (total no. 252). A subset of the first 10 is presented in Table 4.5 which represents the percentage of total variance (derived from Eigenvalue) accounted for each of the PCA and MNF transformed Eagle datasets. For both leaf-on and leaf-off PCA transformed datasets the first principle component accounts for over 97% of the total variance. The first four image bands which exhibited the largest Eigenvalues were examined visually and did not contain noise. The remaining transformed bands exhibited small Eigenvalues and contained a great deal of noise indicating little informational content would be lost if these bands were removed.

For both the leaf-on and leaf-off MNF transformed datasets the percentage of total variance (see Table 4.5) is distributed across the ten bands to a greater degree than for PCA for both the leaf-on and leaf-off datasets. The first 10 MNF bands for both leaf-on and leaf-off datasets contained the highest Eigenvalues accounting for >90% of the total variance. Upon visual inspection these MNF bands did not contain noise. Again the remaining bands had small Eigenvalues and when inspected visually contained a high proportion of noise.

The MNF transformed dataset was selected for future operations as it produced a greater number of usable bands, which were free from noise and captured differing elements within the hyperspectral data to a greater degree than PCA.

Table 4.5– A comparison of Eigen values from leaf-on and leaf-off PCA and MNF transforms for the first 10 bands

	Leaf-on PCA	Leaf-off PCA		Leaf-on MNF	Leaf-off MNF
PCA/MNF band	% of total variance	% of total variance		% of total variance	% of total variance
<i>1</i>	98.38	97.20		36.06	50.17
<i>2</i>	1.29	2.33		19.60	12.72
<i>3</i>	0.19	0.31		8.73	8.49
<i>4</i>	0.06	0.09		6.49	5.63
<i>5</i>	0.03	0.02		4.97	4.78
<i>6</i>	0.02	0.02		4.64	3.56
<i>7</i>	0.01	0.01		4.08	2.53
<i>8</i>	0.01	0.01		3.53	2.35
<i>9</i>	0.01	0.01		2.38	1.80
<i>10</i>	0.01	0.00		2.07	1.71

4.2.2 Spectral indices

Vegetation Indices (VI) are produced from reflectance measurement combinations of several spectral values (between 400 nm to 2500 nm), which are calculated in a manner designed to yield a single value that indicates, for example, the amount of biomass, vegetative vigour or water content (Campbell and Wynne 2011, ENVI-online-help, 2005). Although vegetative foliage, such as leaves, needles and other green materials will often look very similar in the visible wavelengths, they will vary greatly in shape and chemical composition. Such indices may yield insights into the structural properties and condition of the vegetation.

Within ENVI 4.7 there are 21 possible vegetation indices which could be applied to the Eagle dataset. There were eight additional indices available within ENVI software, but these cannot be applied due to specific wavelength requirements (i.e. > 1000nm); for example the assessment of canopy nitrogen. The 21 used indices correspond to five broad groups of indices: (i) Greenness – (broadband and narrowband variants) designed to measure the overall quantity and vigour of vegetation within each pixel; (ii) light use efficiency – designed to estimate a measure of the efficiency with which vegetation is able to use incident light for photosynthesis; (iii) dry or senescent carbon – designed to provide an estimate of the

amount of carbon in its dry states; (iv) leaf pigments – designed to provide an estimate of the stress related pigments present in vegetation (e.g. carotenoids); and (v) canopy water content – designed to provide an estimate of the amount of water held in the foliage canopy. It should be noted that broadband greenness indices are not designed for use with the high spectral resolution hyperspectral datasets, while narrowband greenness indices are.

Within ENVI each of the indices were calculated for both the 2010 leaf-on and leaf-off Eagle hyperspectral data. Eight of the 21 indices did not function as intended and were removed from the analysis, examples include the enhanced vegetation index (broadband greenness), the plant senescence reflectance index (dry and senescent carbon) and the Cartnoid reflectance index 2 (leaf pigments). Appendix B section B.1 summarises the indices attempted in the analysis which functioned as intended. The index images were then exported to ArcMap in GeoTiff format. Zonal statistics functions were then used to extract raster image statistics for each of the field plot polygon extents in addition to a 30x30m grid produced using GME tools (Hawthorne, 2012) (<http://www.spatial ecology.com/gme/>). The summary statistics of maximum, minimum, range, mean, standard deviation and sum were produced for each index within each 30x30m grid-cell, as in Hyde et al. (2006). Table 4.6 summarises the spectral indices which were successfully implemented. A total of 149 metrics were calculated for each cell. A full listing of metrics output for each 30x30m grid cell is available in Appendix B section B.2.

4.2.3 Object-based tree species classification using hyperspectral data

Vegetation interacts with solar radiation differently from other natural materials, such as soils and water bodies. Different plant materials will interact with solar radiation through absorption and reflection in different wavelengths. The different chemical components will be expressed in the reflected optical spectrum from 400 to 2500nm, with spectrally distinct and often overlapping reflectance behaviours. It is possible to classify different landcover materials according to the spectral profile of their image pixels or groups of pixels.

Table 4.6 – A summary list of successfully applied spectral indices using ENVI software

No.	Input dataset:	Index group:	Index name:
1	Leaf-off Eagle	Broadband greenness	NDVI – normalised difference vegetation index
2	Leaf- off Eagle	Broadband greenness	SRI – simple ratio index
3	Leaf- off Eagle	Broadband greenness	ARVI – atmospherically resistant vegetation index
4	Leaf- off Eagle	Narrowband greenness	RENDVI – red edge normalised difference vegetation index
5	Leaf- off Eagle	Narrowband greenness	MRESRI – modified red edge simple ratio index
6	Leaf- off Eagle	Narrowband greenness	MRENDVI – modified red edge normalised difference vegetation index
7	Leaf- off Eagle	Narrowband greenness	VREI – Vogelman red edge index 1
8	Leaf- off Eagle	Light use efficiency	SIPI – structure insensitive pigment index
9	Leaf- off Eagle	Light use efficiency	RGRI – red green ratio index
10	Leaf- off Eagle	Light use efficiency	PRI – photochemical reflectance index
11	Leaf- off Eagle	Leaf pigments	ARI – anthocyanin reflectance index 2
12	Leaf- off Eagle	Leaf pigments	CRI – carotenoid reflectance index 2
13	Leaf- off Eagle	Canopy water content	WBI – water band index
14	Leaf- on Eagle	Broadband greenness	NDVI – normalised difference vegetation Index
15	Leaf- on Eagle	Broadband greenness	SRI – simple ratio index
16	Leaf- on Eagle	Broadband greenness	ARVI – atmospherically resistant vegetation index
17	Leaf- on Eagle	Narrowband greenness	RENDVI – red edge normalised difference vegetation index
18	Leaf- on Eagle	Narrowband greenness	MRESRI – modified red edge simple ratio index
19	Leaf- on Eagle	Narrowband greenness	MRENDVI – modified red edge normalised difference vegetation index
20	Leaf- on Eagle	Narrowband greenness	VREI – Vogelman red edge index 1
21	Leaf- on Eagle	Light use efficiency	SIPI – structure insensitive pigment index
22	Leaf- on Eagle	Light use efficiency	RGRI – red green ratio index
23	Leaf- on Eagle	Light use efficiency	PRI – photochemical reflectance index
24	Leaf- on Eagle	Leaf pigments	ARI – anthocyanin reflectance index 2
25	Leaf- on Eagle	Canopy water content	WBI – water band index

Initially two conventional classification techniques were attempted:(i) maximum likelihood supervised and (ii) IsoData unsupervised within ENVI 4.7 utilising both the entire combined leaf-on and leaf-off datasets and the combined dimensionally reduced datasets. The pixel-based classifiers took a great deal of time to execute and performed poorly when compared with field data. Multiple classes were often found within individual tree crowns, because of differences in illumination and shadow for different parts of the tree crown.

In order to mitigate some of these issues, an approach for tree species classification was employed for this study utilising an object-based method. The object-based method extracts information from images using a hierarchy of image objects, forming groups of pixels for analysis. The approach was developed using the eCognition Developer 8.7 software (Trimble Navigation Ltd., 2012). Conventional segmentation algorithms typically group pixels based upon some degree of similarity between neighbouring groups of pixels. For the changing elevation of a tree crown however this type of segmentation is not applicable. The goal was

to perform a hierarchical classification on automatically delineated individual tree crowns (ITC) of varying dimension in the upper strata of the forest and to classify their species type. A crown-centroid detection approach was performed; this involves the identification of local intensity maxima (bright or high points) and the mapping of crown boundaries by expanding to local minima (Bunting and Lucas, 2006).

4.2.3.1 Development of the eCognition workflow

The process can be divided into four main areas, these are: (i) the construction of a forest area mask; (ii) the delineation of individual tree crown objects; (iii) the broad classification of coniferous and deciduous areas; and (iv) the more specific classification of tree species types.

A combination of leaf-on and leaf-off geo-corrected hyperspectral MNF image-bands and a leaf-on DR LiDAR-derived canopy height model (CHM) (this dataset is defined in section 4.3.3) were added together into an image stack and used for this approach. Table 4.7 summarises all these input datasets. Those hyperspectral MNF image-bands which were chosen contained the most distinction between ground cover types and contained the fewest artefacts.

The object-based classification algorithms were developed in the ‘Cognition Network Language’ (CNL). CNL is a modular programming language allowing typical programming tasks such as branching, looping and variable definition (Tiede et al., 2006). Thus, these modular algorithms can be combined to form a complete ‘ruleware’ for automated information extraction. A workflow was thus devised for the identification of the forest area, the creation of ITC segments and species classification of each object within a class hierarchy.

A class hierarchy is defined broadly as a series of interrelated classes, which form a series of parent and child classes. Only specific child classes can be selected based upon the parent class.

Table 4.7 – Hyperspectral data derived image input list

Dataset Name:	Description:
LiDAR CHM (leaf-on)	LiDAR derived CHM (leaf-on). 1.1x1.1m pixel size
Leaf-on MNF 1	Hyperspectral data derived MNF band 1 from leaf-on eagle data
Leaf-on MNF 2	Hyperspectral data derived MNF band 2 from leaf-on eagle data
Leaf-on MNF 3	Hyperspectral data derived MNF band 3 from leaf-on eagle data
Leaf-on MNF 4	Hyperspectral data derived MNF band 4 from leaf-on eagle data
Leaf-on MNF 5	Hyperspectral data derived MNF band 5 from leaf-on eagle data
Leaf-on MNF 6	Hyperspectral data derived MNF band 6 from leaf-on eagle data
Leaf-off MNF 1	Hyperspectral data derived MNF band 1 from leaf-off eagle data
Leaf-off MNF 2	Hyperspectral data derived MNF band 2 from leaf-off eagle data
Leaf-off MNF 3	Hyperspectral data derived MNF band 3 from leaf-off eagle data
Leaf-off MNF 4	Hyperspectral data derived MNF band 4 from leaf-off eagle data
Leaf-off MNF 5	Hyperspectral data derived MNF band 5 from leaf-off eagle data
Leaf-off MNF 6	Hyperspectral data derived MNF band 6 from leaf-off eagle data

4.2.3.2 Creation of the forest mask

The forest mask process used only the CHM raster layer as input. This process began with the use of a contrast-split segmentation, where the raster was segmented based upon the contrast between high and low regions (Trimble Navigation LTD., 2012). Matthews and Mackie (2006) broadly define saplings as tree species which are below 1.3m in height; heights above this threshold are considered as trees. Thus, raster grid-cell values above a height threshold value of 1.3m were considered forest. Initial segments were created around these regions classifying high ($>1.3\text{m}$) areas as forest and low ($\leq 1.3\text{m}$) areas as non-forest.

What followed was an additional optional step to classify any segments which remained unclassified based upon the mean height values of the object. The final operations were to merge all forest or non-forest segments. This step created the first level of the class hierarchy, forested and non-forested areas.

4.2.3.3 Individual tree crown delineation

The method used to segment the data into ITC objects is a modification of the approach presented in Bunting and Lucas (2006), where the crown-centroids were detected through the use of a LiDAR-derived CHM instead of spectral data. Tree tops or ‘local maxima’ were detected by a moving search window, once found these maxima cells were reclassified to crown seeds. Further iterative steps were then implemented for growing ITC objects around these seed points into grid-cells of lower elevation using a progressive set of rules to delineate the crowns/clusters. A description of the process employed within this project follows.

The CHM raster layer height values were used as input for the ITC delineation procedure. The first step was to identify any isolated tree crowns from the application of the forest mask; i.e. these trees could exist separately from larger forested areas and would be surrounded by non-forest objects. This was performed by identifying these forest objects based on size and shape parameters of the object segment (e.g. area, elliptic fit, length/width ratios, and roundness). If the tree crown shape and size criteria were met these objects were given the class of ‘tree crown’, and not considered in further delineation steps.

The larger forest segments, or clusters of trees, were visited and processed iteratively. The first process within each iteration was to perform a chessboard segmentation, effectively separating that cluster region into its individual raster grid-cells (1x1m), or rather each raster grid-cell in a forest classified area had a corresponding object. An iterative procedure was then implemented to delineate ITC objects. In theory, tree crowns form ‘dome’ shapes within the CHM raster surface, where the higher (brighter) dome tops correspond with the crown tops. Therefore by identifying these tops as seeds and expanding the seeds into neighbouring areas of lower height the ITCs were delineated. ITC object growing could also halt in a particular direction if it came into contact with other seed or crown object boundaries. The ‘find local extrema’ algorithm was used to find the highest points within a mobile initial search window size with a radius of 5m. What followed was an iterative expansion of tree crown objects from the seed points, where 1x1m objects bordering the seed point of the same or lower height were merged, and the process repeated until height values bordering the object increased or the object boundaries came into contact with other ITC objects or non-forested areas. This process was applied to each of the forest cluster objects.

The final stage was to merge any remaining unclassified forest or crown objects (that were left where no crown seeds have been identified) and then re-apply the isolated tree crown identification procedure using the aforementioned criteria (area, elliptic fit, length/width ratios and roundness). Those objects which remained unidentified were assessed individually by the user.

The result of this processing is a map of overstorey tree crown locations, see Figure 4.7. It should be noted that this was a very computationally intensive process, taking on average one hour for a single 1x1km area (at a spatial resolution of 1m).

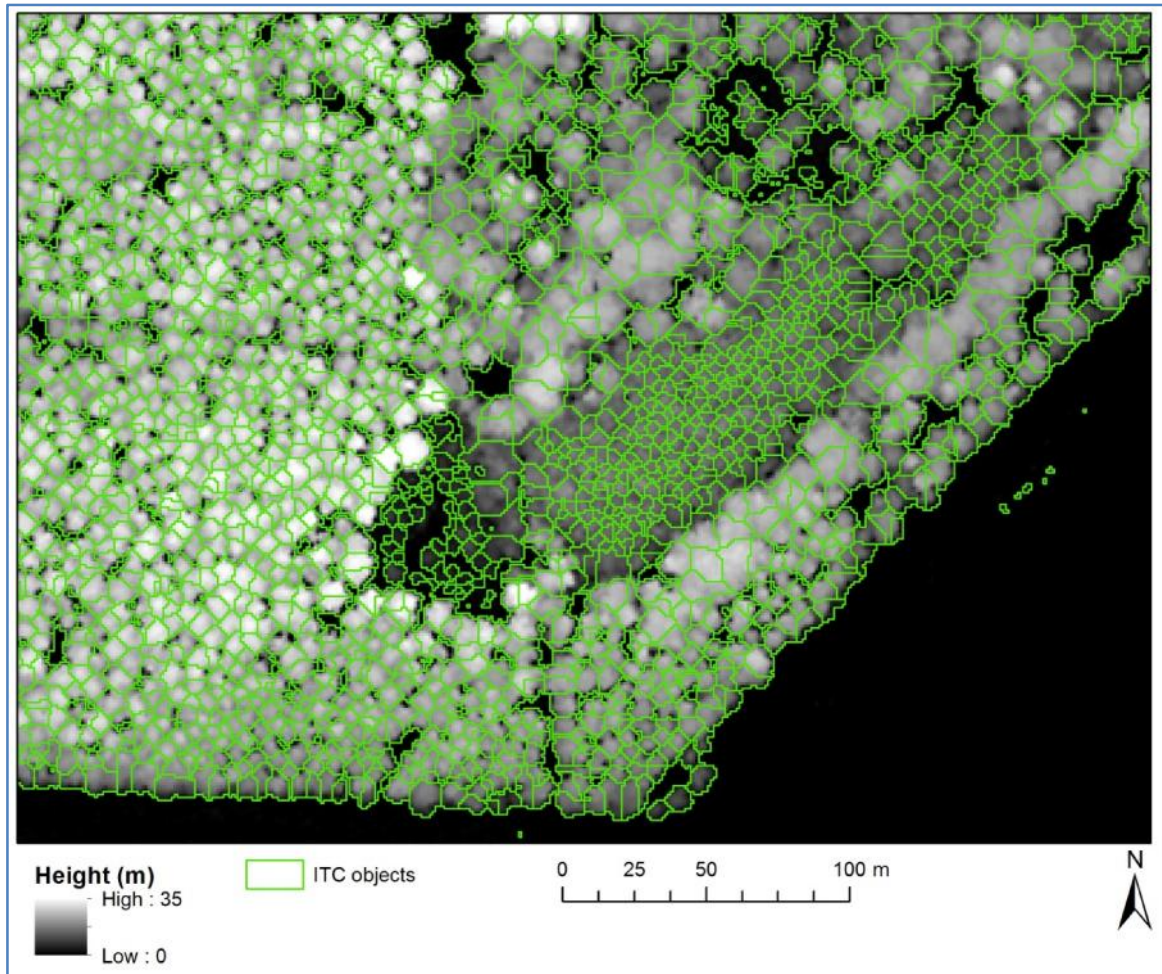


Figure 4.7 – An example of the ITC segments created through the crown-centroid detection approach, overlaid on the CHM.

4.2.3.4 Broad tree species classification

A broad classification was then applied to the ITC segments to distinguish deciduous and coniferous species types. This was accomplished through a membership function applied to the mean of the groups of pixels intersecting each ITC object, which was applied to only forest classified objects. Values derived from a single hyperspectral image, MNF3 (leaf-on), was utilised to determine which areas belonged to the different broad species types. For the case of the MNF 3 (leaf-on) image, it was found that it best exemplified the difference between coniferous and deciduous species. Thus, the average pixel value was calculated for those pixels in the MNF3 (leaf-on) image intersecting each ITC segment. The classification was based upon membership function thresholds, where higher values (2 to 50) were classified as a coniferous ITC, whereas lower values (-40 to 2) were classified as deciduous.

An additional class was added at this stage to encompass man-made structures and potential edge effects. The former related to the small number of houses located in the field plot

region, whereas the latter related to the slight misalignment of LiDAR derived ITC objects and the hyperspectral image, often at the edge between forest and open areas, where non-tree pixels could skew the results of the classification. Both of these features could be determined by high pixel values (>50) in the leaf-on MNF3 image. These objects were then removed from the classification process.

This step formed the second level of the class hierarchy where forested areas became one of the three classes, coniferous, deciduous or man-made.

4.2.3.5 In-depth tree species classification

The next step was to add another level in the class hierarchy; a more detailed classification of tree species types within the existing coniferous and deciduous class contexts. Those species types the user would expect to occur within the study area, according to fieldwork and FC inventory data, were entered into the appropriate class hierarchy. For example the deciduous class was expanded to include many sub-classes such as oak, beech, silver birch, sweet chestnut, and holly. Likewise the coniferous class could be broken down into sub-classes such as scots pine, corsican pine, douglas fir, or norway spruce. All input hyperspectral MNF rasters (leaf-on and leaf-off) including those raster layers derived from LiDAR data could be used as an input to the classification process.

Owing to the presence of plantation woodland and the age differences in the tree species between stands an additional step was added to the classification: a separation of older and younger trees. Difference in ages within a tree species group could potentially result in a different hyperspectral signature within the MNF-bands. An additional classification step was applied to both the coniferous and deciduous classified ITC objects. A ‘height filter’ was used to classify those objects above a certain height threshold as ‘mature’ woodland (CHM $>15\text{m}$ in height), whereas values below this threshold were classified as ‘immature’ woodland (CHM $<15\text{m}$ in height). The choice of the threshold was determined by information provided in the Forestry Commission inventory data and unsatisfactory classification results whilst experimenting with the method. Recently planted tree species were typically below 15m in height and some of which exhibited different pixel values to those of the same species in other compartments.

A number of studies (e.g. Gougeon and Leckie, 2003; Leckie et al., 2005) have extracted tree spectra from the mean-lit (sunlit) sections of the proportions of the crown rather than individual pixels. MNF image values cannot be related to any spectral measurements because of how they were calculated. Information held within the MNF images is sufficient for classification purposes however (Onojeghuo and Blackburn, 2011). The dimensionally reduced Eagle datasets were then used to identify young and mature tree species types.

Each of the input images was interrogated by calculating the mean or maximum value of the image layer's pixels which intersect each ITC object, creating a value for each. Classifying tree species from ITC objects is a complicated task due to the variability of each of the hyperspectral input data for each object. A number of membership functions were developed manually for each tree species class. Each species potentially utilise a number of input images, combinations of functions based on the relationships between that ITC object and its neighbours. Logical operators (e.g. AND, OR, NOT) were also used to account for conflicting or consistent features between different class membership functions. A full list of all classes and membership functions is listed in Table 4.8.

The final stage to complete within eCognition was to export the classified ITC layer, using the 'export as vector' function. The output file was set as an ESRI format shapefile, where each ITC object was converted to a polygon, and the species type added as an attribute in the linked database table. A total of 28 tree species classes were developed.

Table 4.8 – Class list and membership functions

Class name	Class hierarchy (parent classes)	Membership function 1	Membership function 2	Membership function 3
Forest	-	N/A – assigned by segmentation	-	-
Non-forest	-	N/A – assigned by segmentation	-	-
Crown	<i>Forest</i>	N/A – assigned height maxima search and region growing	-	-
Coniferous	<i>Crown</i>	Difference in Leaf-on MNF image 3 (values: $-40 > x < 2$)	-	-
Deciduous	<i>Crown</i>	Difference in Leaf-on MNF image 3 (values: $2 > x < 50$)	-	-
Manmade structures and edge effects	<i>Crown</i>	Difference in Leaf-on MNF image 3 (values: $50 > x$)	-	-
Immature Coniferous	<i>Coniferous</i>	Maximum CHM value within object (values: $1.3 > x < 15$)	-	-
Mature Coniferous	<i>Coniferous</i>	Maximum CHM value within object (values: $15 > x < 50$)	-	-
Immature Deciduous	<i>Deciduous</i>	Maximum CHM value within object (values: $1.3 > x < 15$)	-	-
Mature Deciduous	<i>Deciduous</i>	Maximum CHM value within object (values: $15 > x < 50$)	-	-
(Young) corsicanpine (<i>Pinusnigra</i>)	<i>Immature Coniferous</i>	Mean leaf-on MNF 2 (values $-6 > x < -1$)	-	-
(Young) douglasfir (<i>Pseudotsugamenziesii</i>)	<i>Immature Coniferous</i>	Mean leaf-on MNF 2 (values $-13 > x < -6$)	<NOT>Mean leaf-off MNF 4 (values $-16 > x < 0$)	<NOT>Mean leaf-on MNF 3 (values $-11 > x < 11$)
(Young) grand fir (<i>Abiesgrandis</i>)	<i>Immature Coniferous</i>	Mean leaf-on MNF 2 (values $-15 > x < -6$)	MeanMNF 3 (values $-1 > x < 11$)	MeanMNF 4 (values $-4 > x < 7$)
(Young) japaneselarch (<i>Larixkaempferi</i>)	<i>Immature Coniferous</i>	Mean leaf-off MNF 4 (values $-16 > x < 0$)	-	-
(Young) hybrid larch (<i>Larixeurolepis</i>)	<i>Immature Coniferous</i>	Mean leaf-on MNF 5 (values $2 > x < 7$)	-	-
(Young) norwayspruce (<i>Piceaabies</i>)	<i>Immature Coniferous</i>	Mean leaf-on MNF 4 (values $-1 > x < 7$)	<NOT>Mean leaf-on MNF 3 (values $-25 > x < -14$)	<NOT>Mean leaf-on MNF 3 (values $-1 > x < 11$)
(Young)scots pine (<i>Pinussylvestris</i>)	<i>Immature Coniferous</i>	Mean leaf-on MNF 2 (values $-1 > x < 6$)	-	-
(Young) western hemlock (<i>Tsugaheterophylla</i>)	<i>Immature Coniferous</i>	Mean leaf-on MNF 3 (values $-25 > x < -14$)	-	-

Table 4.8 (continued)

coast redwood (<i>Sequoia sempervirens</i>)	<i>Mature Coniferous</i>	Mean leaf-on MNF 2 (values $-15 > x < -6$)	Mean leaf-on MNF 3 (values $-1 > x < 11$)	Mean leaf-on MNF 4 (values $-4 > x < 7$)
corsicanpine	<i>Mature Coniferous</i>	Mean leaf-on MNF 2 (values $-1 > x < 20$)	-	-
douglasfir	<i>Mature Coniferous</i>	Mean leaf-on MNF 2 (values $-8 > x < -2$)	Mean leaf-on MNF 3 (values $-7 > x < 0$)	-
grand fir	<i>Mature Coniferous</i>	Mean leaf-on MNF 2 (values $-15 > x < -6$)	MeanMNF 3 (values $-1 > x < 11$)	MeanMNF 4 (values $-4 > x < 7$)
hybrid larch	<i>Mature Coniferous</i>	Mean leaf-on MNF 5 (values $2 > x < 7$)	-	-
japaneselarch	<i>Mature Coniferous</i>	Mean leaf-off MNF 4 (values $-16 > x < 0$)	-	-
norwayspruce	<i>Mature Coniferous</i>	Mean leaf-on MNF 4 (values $-1 > x < 7$)	<NOT>Mean leaf-on MNF 3 (values $-25 > x < -14$)	<NOT>Mean leaf- on MNF 3 (values $-1 > x < 11$)
scots pine	<i>Immature Coniferous</i>	Mean leaf-on MNF 2 (values $-1 > x < 6$)	-	-
lawsons cypress (<i>Chamaecyparisl awsoniana</i>)	<i>Mature Coniferous</i>	Mean leaf-off MNF 4 (values $7 > x < 13$)	-	-
western hemlock	<i>Mature Coniferous</i>	Mean leaf-on MNF 3 (values $-25 > x < -4$)	-	-
(Young) common alder (<i>Alnusglutinosa</i>)	<i>Immature Deciduous</i>	Mean leaf-on MNF 4 (values $-13 > x < -6$)	-	-
(Young) oak (<i>Quercusrobur</i>)	<i>Immature Deciduous</i>	Mean leaf-on MNF 4 (values $-5 > x < 2$)	-	-
(Young) beech (<i>Fagussylvatica</i>)	<i>Immature Deciduous</i>	Mean leaf-on MNF 4 (values $2 > x < 25$)	-	-
(Young) silver birch (<i>Betulapendula</i>)	<i>Immature Deciduous</i>	Mean leaf-on MNF 4 (values $-20 > x < -5$)	-	-
(Young) sweet chestnut (<i>Castaneasativa</i>)	<i>Immature Deciduous</i>	Maximum pixel values of leaf-on MNF 3 (Values from 24 to 45)	-	-
common alder	<i>Mature Deciduous</i>	Mean leaf-on MNF 4 (values $-13 > x < -6$)	-	-
oak	<i>Mature Deciduous</i>	Mean leaf-on MNF 3 (values $1 > x < 20$)	<AND>Mean leaf-on MNF 4 (values $-7 > x < 2$)	-
beech	<i>Mature Deciduous</i>	Mean leaf-on MNF 4 (values $2 > x < 25$)	-	-
silver birch	<i>Mature Deciduous</i>	Mean leaf-on MNF 4 (values $-20 > x < -7$)	-	-
sweet chestnut	<i>Mature Deciduous</i>	Maximum pixel values of leaf-on MNF 3 (Values from 24 to 45)	-	-

4.2.3.6 Classification of 30x30m summary metrics

Area-based summaries of the classified ITC-objects were required for input into statistical modelling for the estimation of forest attributes. The resulting classified ITC map was then exported as a shapefile for use in ArcMap. Point-centroids were extracted from the ITC polygons using the GME software. Summary metrics were extracted for the 30x30m field-plot extent shapefiles (generated from the coordinates recorded in the two fieldwork operations) and for a regularly spaced grid of 30x30m shapefiles produced by the GME software for the whole study site. A spatial join operation was then performed with the field-plot location shapefiles, and separately, for the whole study site.

Area based metrics were extracted and/or calculated for the 30x30m extents using custom R code (<http://cran.r-project.org/>), documented in Appendix B section B.3. This R script calculated the number of ITC objects, number of species and number of native species. A Shannon and Simpson index was calculated using the R package: Vegan (<http://cran.r-project.org/web/packages/vegan/index.html>) (Oksanen et al., 2012). A total of eight output metrics were derived, these were: (i) number of ITC objects which intersected the grid cell; (ii) number of native tree species classified objects which intersected the grid cell; (iii) number of tree species encountered; (iv) number of native tree species encountered (i.e. of the species: oak, beech, silver birch, scots pine, or common alder); (v) Shannon Index calculated from counts and species within the cell; (vi) Simpson Index calculated from counts and species within the cell; (vii) Evenness index, derived from the Shannon Index; (viii) a count of the ITC objects of the same species with the largest population relative to the others in the 30x30m extent; and (ix) a count of the ITC objects of the same species with the lowest population relative to the others in the 30x30m extent. The initial six metrics were extracted for the field plot areas and used for statistical analysis. The remaining two were required for use as input for the Complex Stand Diversity Index (Jaehne and Dohrenbusch, 1997 in Vorčák et al., 2006).

4.3 Discrete-return LiDAR processing

This section describes the pre-processing steps required in order to allow analysis with the LiDAR datasets.

The 2010 LAS files were delivered by the ARSF in a format in which the following processing steps could be directly applied. It should be noted that for each of the pre-processing steps listed in sections 4.3.1 – 4.3.3 both the leaf-on and leaf-off DR LiDAR datasets were processed independently of one another. Only leaf-on LiDAR data was utilised in the ITC delineation however.

The initial quality assessment provided by the ARSF prior to delivery of the datasets indicates the geometric accuracy of the LiDAR data agrees with the Ordnance Survey vectors on average to within 0-1 metres. The geometric accuracy for the scanner is stated as a vertical nominal accuracy of 0.05-0.10m, and a horizontal accuracy of between 0.13-0.61m (Leica Geosystems, 2003).

4.3.1 Filtering ground and non-ground points

LAS file processing tasks, such as noise removal from the delivered files and sub-setting of areas were performed using the ‘LAStools’ software (<http://www.cs.unc.edu/~isenburg/lastools/>) (Isenburg, 2013). The LAS files were delivered with a basic classification already applied by the Terrascan software (<http://www.terrasolid.fi>). This classification identified probable ‘noise’ points, which could be removed from the dataset completely. In order to remove noise points the ‘lasclip’ tool provided in LAStools was used. This set of tools allows for the removal of noisy areas completely, in addition to the removal of noise identified above or below certain limits in all three dimensions. Figure 4.8 illustrates the effects of clouds, some distance above the ground, upon the data, which must be removed. Fortunately all these erroneous areas could be removed from the point data. Point data from adjacent flightlines were used as a substitute (i.e. to fill the gaps). Once the point cloud was modified to remove these erroneous effects the subsets were sequentially merged back into a single LAS file containing all the point data for the study area using ‘LASmerge’ (LAStools).

As Sithole and Vosselman (2004) indicate, the filtering of bare-earth airborne LiDAR points is a crucial procedure for LiDAR data processing. This particular step can take up to 60 to 80 per-cent of the processing time (Chen et al. 2007b). The basic principle of surface-based methods is to create a surface with a corresponding buffer zone above it, where the buffer zone defines the region in 3D space where terrain points are expected to reside (Sithole and Vosselman 2004). The method creates a surface approximating the bare-earth elevations.

Filtering the LiDAR point cloud into ground and non-ground returns is required in order to generate bare-earth data and perform further analysis, such as deriving height information for trees and buildings. The RSC LAStools software (<http://code.google.com/p/rsclastools/>) was utilised to filter the DR LiDAR datasets. This software was provided by John Armston, from the University of Queensland, Australia (Armston, 2011). The program was written in Interactive-Data-Language (IDL) (ITT Visual Information Solutions) and runs through the IDL Virtual Machine with a (32-bit) Windows XP desktop computer. The filtering method used in the software is outlined in Zhang et al.(2003), and utilises a progressive morphological filter. The algorithm classifies non-ground and ground LiDAR measurements by passing a search window across the extent of the LiDAR point cloud. It has been found that fixed window sizes can encounter difficulty in detecting all non-ground objects. This problem can be solved by increasing the window sizes of morphological filters gradually through a number of iterations. Using elevation and difference thresholds, the measurements of buildings and vegetation are removed while ground vegetation is preserved. Figure 4.9 illustrates a classified point cloud and a simple representation of a ground model. The initial settings for these filtering procedures are not automatic, and it is necessary for a number of attempts in order to produce an optimal classification. Owing to the often changing forest structure located in each of the flight strips, increased search window sizes were used for all filtering operations. The window size was set to 25m in order to identify the lowest point in often densely vegetated areas. The filtering technique process is not automatic, filter parameters needed tuning and the process repeated in order to get the best results.

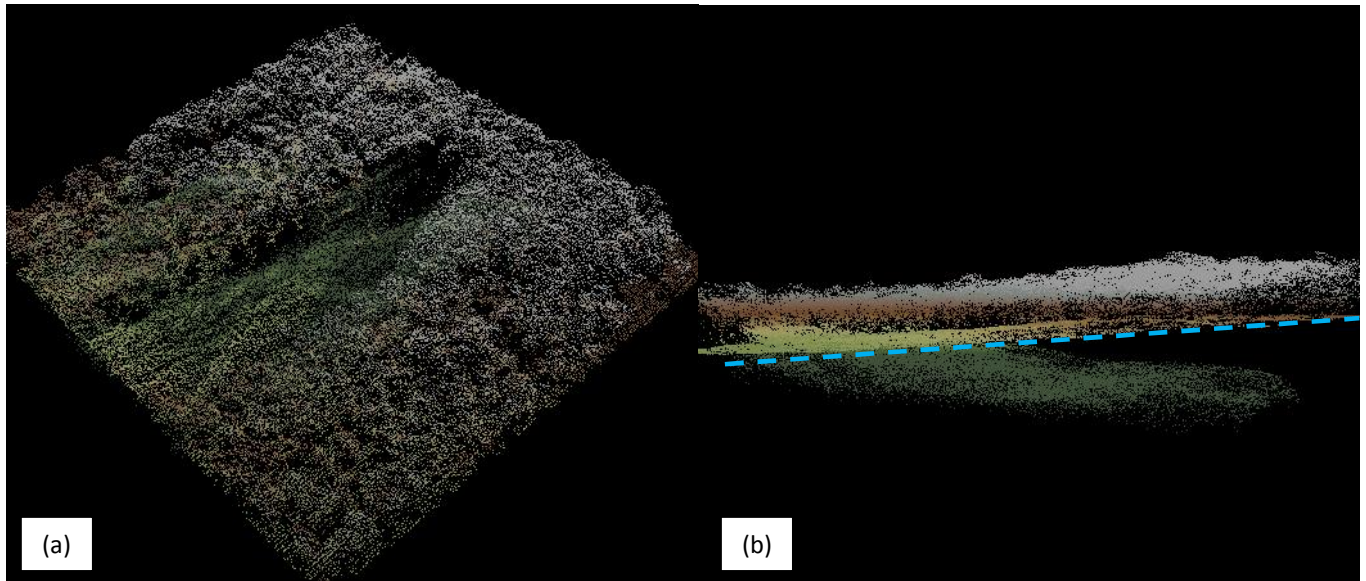


Figure 4.8 – Cloud within the flightline caused erroneous measurements in the data. As seen in the left image (a) where data about the tree canopy is lost. The right image (b) shows the extent of the errors below what is supposed to be the ground surface (blue line). This area was removed from the analysis. Screenshots taken in the Fugro Viewer software (<http://www.fugroviewer.com/>).

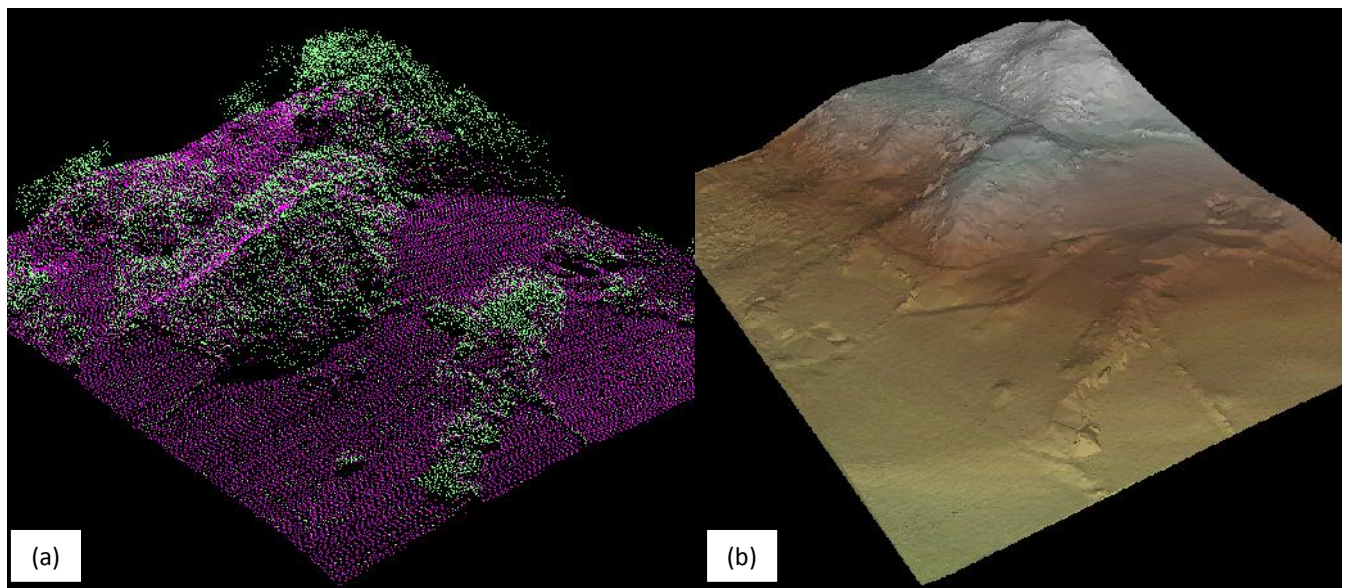


Figure 4.9 – (a) Gives an example of a classified point cloud where green is classified as non-ground (mostly vegetation within this scene) and purple as ground. (b) A Triangular-Irregular-Network (TIN) of the ground points. Screenshots taken in the Fugro Viewer software (<http://www.fugroviewer.com/>).

4.3.2 Ground separation and normalisation

In order to remove the effects of terrain variation upon forest tree height measurements it is necessary to interpolate a DTM from the classified ground points. Once interpolated the DTM is used to subtract the elevations of the terrain from the remaining points within the LAS file, normalising the dataset to above-ground height. With RSC LAStools, the DTM was interpolated using the Natural Neighbour interpolation method. The DTM produced from these operations is then used to remove the effects of terrain from the non-ground classified points, effectively subtracting the terrain elevations from the rest of the point cloud, as illustrated in Figure 4.10.

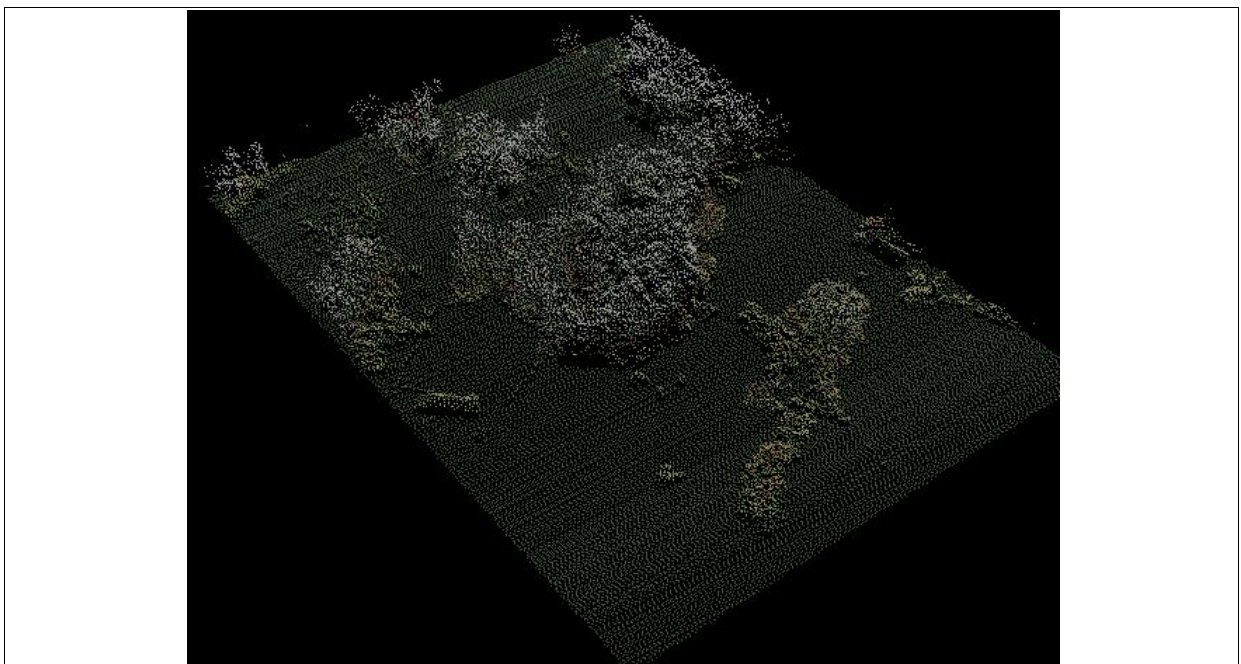


Figure 4.10 – The same area as in figure 4.8, with the effects of the underlying terrain removed

4.3.3 Extraction of DR LiDAR structure related metrics

It was necessary to extract area-based metrics for the same spatial extent of the field data collection, in this case 30x30m. Metrics were extracted for all of the field-plot locations and separately for the 30x30m grid shapefile of the whole study site. To reiterate the GME software was used to generate a coverage of 30x30m grid cell shapefiles for the entirety of the New Forest study site.

The RSC LAStools software was capable of producing the common rasterised surface layers, such as the DTM, DSM, intensity image (from first returns only), and CHM. Raster layers can be generated for various attributes for both ground and non-ground points. ASCII summary statistics were calculated, including metrics such as canopy openness, height percentiles, mean, variance, skewness and kurtosis. Summary statistics were generated for (i) all the points, (ii) those points classified as ground and (iii) those points classified as non-ground. Table 4.9 reports all the extracted indicator metrics from this software for both leaf-on and leaf-off datasets. All of the extracted metrics have been used in previous research as a way of classifying forest structure (Falkowski et al. 2009a; Hudak et al. 2008). Each of these metrics was calculated for a defined output extent of 30x30m. Table 4.10 describes mathematically the necessary calculations. A full listing of extracted metrics is given in Appendix C sections C.1 and C.2. The next stage in the research was to perform statistical analysis utilising these area-based metrics and their relationships to the metrics collected in the field. A total of 187 metrics were extracted for each field plot extent, this included leaf-on and leaf-off datasets.

RSC LAStools can produce a number of additional statistics which were not included in this analysis. Point density metrics were not used due to the differences between areas where flightlines overlap and do not overlap, leading to obvious disparities between areas. LiDAR intensity was uncalibrated as its values could be altered during the acquisition flight (Lindberg et al., 2012). Hölje and Pfeifer (2007) also state a number of problems if using the original intensity values relating to the influences of topographic and atmospheric effects which would influence the backscatter of the emitted laser pulse. Thus, metrics relating to the maximum, minimum and range of intensity values were removed. Instead only statistics regarding the distribution of intensity values, such as skewness and kurtosis, were employed.

In addition a CHM raster layer was produced in order to delineate ITC objects in eCognition as reported in section 4.2.3. The spatial resolution of the LiDAR layer was produced to be 1.1x1.1m (i.e. the same as the Eagle resolution), and the raster grid was aligned to that of the hyperspectral data.

Table 4.9 – Extracted DR LiDAR metrics used in this study

Ground and above-ground height metrics:	Intensity:
Ground terrain metrics (slope/roughness)	-
Canopy cover	-
Canopy relief ratio	-
Maximum	-
Mean	Mean
Median	Median
Standard deviation	Standard deviation
Variance	Variance
Absolute deviation	Absolute deviation
Skewness	Skewness
Kurtosis	Kurtosis
Percentiles at 5% intervals (5,10...90,95)	Percentiles at 5% intervals (5,10...90,95)

Table 4.10 – Mathematical descriptions of computed metrics. Metrics are calculated directly from binned point cloud and can be calculated on elevation, heights, intensity and point density. Where x = numeric variable, n = number of observations, μ = mean, σ = standard deviation, and λ = frequency (Evans et al., 2009).

Metric:	Description:
Topographic slope	The topographic slope estimated from a DEM generated from the ground classified points (Burrough and McDonell, 1998).
Topographic roughness	The topographic roughness estimated from a DEM generated from the ground classified point.
Canopy Cover	$CC = \left(\frac{h_{ng}}{h_{all}} \right)$ <p>Where h_{ng} and h_{all} denote the sum total of non-ground returns and the sum of all returns respectively.</p>
Canopy relief ratio	$E = \frac{h_{mean} - h_{min}}{h_{max} - h_{min}}$ <p>Where h_{mean}, h_{min} and h_{max} are the mean, minimum and maximum canopy heights, respectively (Pike and Wilson, 1971).</p>
Minimum	Minimum value (x)
Maximum	Maximum value (x)
Range	[Maximum value (x) - Minimum value (x)]
Arithmetic Mean (μ)	$\mu = \frac{\sum_{i=1}^N x_i}{N}$
Standard Deviation (σ)	$\sigma = \sqrt{\frac{\sum_{i=1}^N x_i^2}{n} - \mu^2}$
Variance (σ^2)	$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n}$
Mean Absolute Deviation from Mean (MADM)	$MADM = \mu \left(\frac{\sum_{i=1}^n x_i - \mu }{n} \right)$
Skewness	$Skewness = \frac{\sum_{i=1}^n (x_i - \mu)^3}{\left(\sum_{i=1}^n (x_i - \mu)^2 \right)^{3/2}}$
Kurtosis	$Kurtosis = \frac{\sum_{i=1}^n (x_i - \mu)^4}{\left(\sum_{i=1}^n (x_i - \mu)^2 \right)^2} - 3$

4.3.4 Individual tree crown delineation

As stated in section 4.2.3 an automated ITC delineation approach was implemented within eCognition. The ITCs were produced primarily for species classification purposes. Additional metrics such as tree crown radius were not readily available and would require additional calculations to be produced and implemented in order to extract them. Another piece of software was available to provide automated ITC delineation which would provide additional metrics regarding the individual tree crowns location, crown horizontal radius, area and geometric volume. Such ITC crown size metrics have been identified as important inputs into predicative models in other studies (e.g. Hyypä et al., 2001; Person et al., 2002; Popescu et al., 2004; Maltamo et al., 2004). The Toolbox for LiDAR Data Filtering and Forest Studies (TIFFS) software (<http://www.globalidar.com>) was utilised to this end. TIFFS is commercial software provided by Globlidar, which utilises the Matlab runtime library.

The identification of individual tree crowns (ITC) was performed using leaf-on LiDAR data only. The above process of filtering ground and non-ground points and normalisation of non-ground returns produced using RSC LAStools could not be utilised within the TIFFS software. Thus it was necessary to repeat the filtering pre-processing steps within the TIFFS software.

The first part of data processing in TIFFS was to tile the raw LiDAR data of all the recorded flight strips, re-ordering the data into grid cells which can combine parts of different flightlines. The filtering method used is that of Chen et al. (2007b) where a mobile filtering window was utilised to identify the minimum elevation. Different mobile filtering search window sizes were tested and a size of 25m diameter was found to be optimal in this context. The lowest elevations recorded were interpolated into a grid using the Nearest Neighbour method. This grid now becomes the Digital Terrain Model (DTM). As with the previous filtering technique the process is not automatic, filter parameters need tuning in order to get the best results. The DTM creation conforms to the definition of Nearest Neighbour interpolation. The DTM produced was subsequently used to normalise the non-ground data points to derive canopy height.

Trees were isolated using a marker-controlled watershed segmentation method as used in Chen et al.(2006). The treetops were detected by searching for local maxima in a Canopy Maxima Model (CMM) with variable window sizes. Window sizes varied depending on the

lower-limit of the prediction intervals of the regression curve between crown size and tree height. Chen et al. (2006) reports the CMM was created to reduce the commission errors of treetop detection. The TIFFS software extracted individual tree height, crown radius and 3D crown area of those trees in the forest overstorey. The extraction technique is based upon the theory presented in Chen et al.(2007a), predicted through a metric termed ‘canopy geometric volume’ (CGV). The CGV is the volume encircled by the outer surface of the crown, which can be derived easily by combining the canopy height model and individual-tree crown map. The default output of this process is in the shapefile format for Esri ArcGIS. Table 4.11 summarises the attributes which can be extracted for each potential tree crown available in the shapefile generated by TIFFS. Figure 4.11 illustrates the estimation of tree crown location (points) and area (polygon) overlaid upon the CHM.

In order to utilise these overstorey canopy metrics for the 30x30m grid based analysis the ITC point metrics were summarised for each of the field-plot location shapefiles and for the 30x30m grid coverage of the whole study site. Spatial join operations within ArcMap 10 were utilised to calculate the average and total values for all ITC objects (trees) located within a cell polygon, metrics included the number of tree points, height, crown radius, horizontal area and CGV. All gridded-ITC summary metrics are listed in Appendix C section C.3. Metrics relating to the number of trees, average height, average and total crown area, average and total crown volume, the average tree nearest neighbour (NN) spacing and standard deviation of the tree NN spacing were all calculated using the statistical package R, and is reported in Appendix C section C.4. The spacing of trees was calculated using the X and Y coordinates of each ITC point intersecting the 30x30m grid-cell polygon, where mean spacing and standard deviation were calculated. These ITC metrics were then added to the 30x30m grid shapefile based on cell ID. Nine summary metrics were produced; these are given in Table 4.12.

The construction of the complex stand diversity index (CSDI) required a number of additional ITC related metrics for its construction, these were the three largest and three smallest tree stem NN distances, in addition to two maximum and two minimum tree crown horizontal diameters within the 30x30m field-plot or grid cell area. These data were extracted using an R-script, which is reported in Appendix C section C.5.

Table 4.11 – All extracted ITC metrics for a single crown derived from TIFFS

Metric:	Metric description:
XY coordinates	The geometric coordinates of the tree crown maxima (assumed centre).
crownRad	Crown radius calculated from tree crown centre to edge of supposed crown (m).
Crown_area*	Crown horizontal area was calculated from the crown radius using the equation for the area of a circle: $A = \pi r^2$ where A is the crown area in metres square, and r is the crown radius.
treeHt	Height of the tree crown maxima point (m)
canopyVol	The canopy geometric volume (GMV) (m ³) (Chen et al., 2007a)

*Derived from data extracted in the TIFFS software

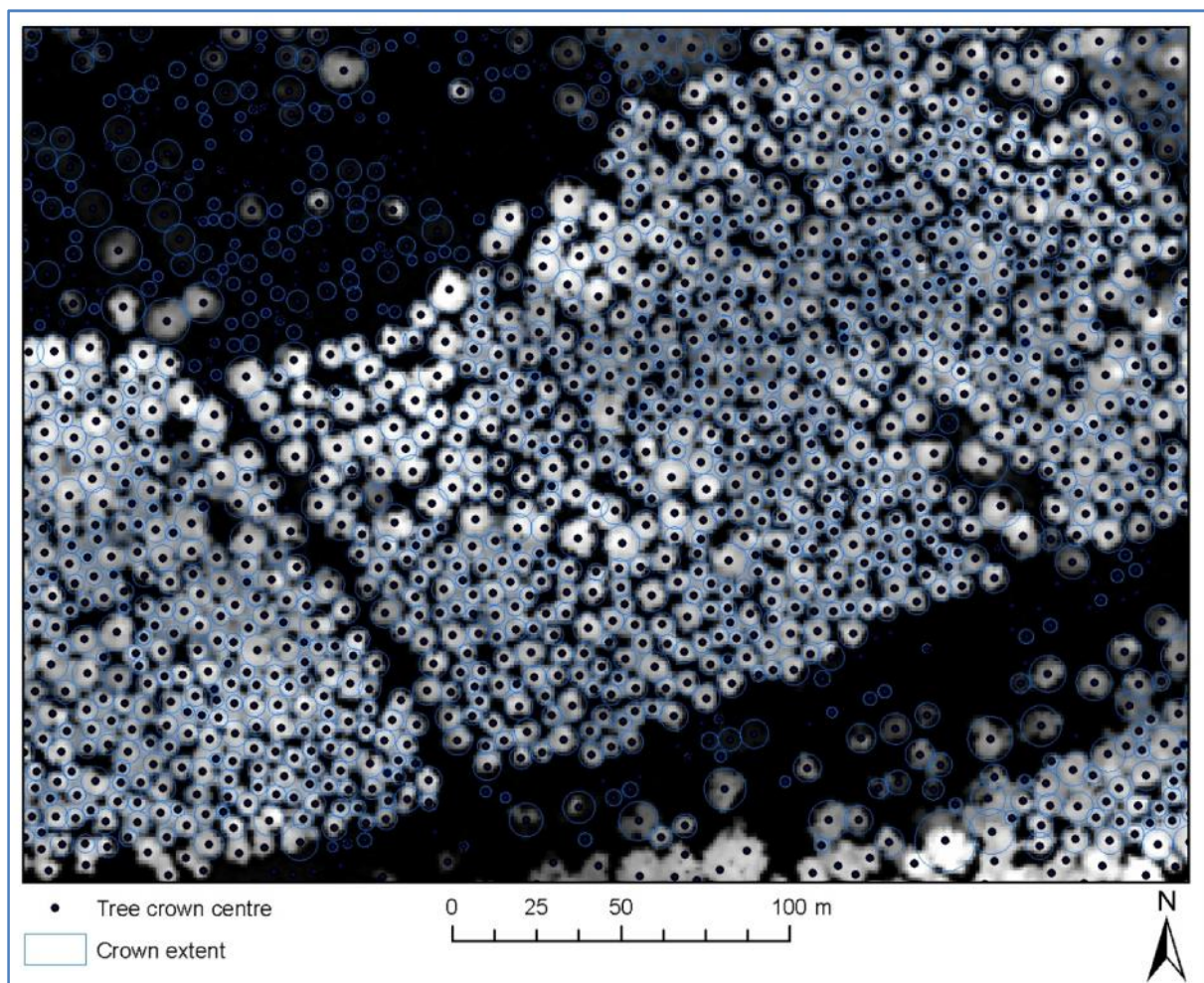


Figure 4.11 – ITC points and polygons generated through TIFFS, shown for conifer stands of plantation forest.

Table 4.12 – ITC area summary values of ITC objects within each 30x30m grid cell

Metric:	Description:
TIF_TreeNo	A count of ITC objects within a given 30x30m area
TIF_Space	The average stem nearest neighbour (NN) distance (m)
TIF_STD	The standard deviation of stemnearest neighbour (NN) distance (m)
TIF_CR	Average crown radius (m)
TIF_m_area	Average crown horizontal area (m ²)
TIF_to_area	Total crown horizontal area (m ²)
TIF_HT	Average tree height (m)
TIF_meanCV	Average crown geometric volume (m ³)
TIF_totalCV	Total crown geometric volume (m ³)

4.4 Full-waveform LiDAR processing

For leaf-on and leaf-off 2010 acquisitions small-footprint full waveform LiDAR was provided by the ARSF. The individual flight strips were delivered in LAS (version 1.3) format, which provides coordinate information of the aircraft, the X, Y and Z location of the first major return for each waveform, and the 256 waveform samples for each laser pulse. At the time of writing, there are few examples of processing software capable of extracting any coordinate data from a small-footprint waveform system, or more specifically this particular file format.

It should be noted that for each of the pre-processing steps listed in sections 4.4.1 – 4.4.2 both the leaf-on and leaf-off FW LiDAR datasets were processed independently of one another.

The use of FW LiDAR posed a number of problems, where one of the major issues was the lack of software being capable of interpreting the unique format of the files and processing the waveform into point information for analysis. One of the possible software was that of ‘orientation and processing of airborne laser scanning data’ (OPALS) (<http://geo.tuwien.ac.at/opals/html/index.html>). Unfortunately at the time of this project OPALS could only operate upon proprietary Riegl data formats for FW LiDAR data. Instead the Sorted Pulse Software Library (SPDlib) (<http://www.spdlib.org/>) software was used, which was developed by Bunting et al. (2011b, 2011a). SPDlib is set of open source software tools for processing laser scanning data.

SPDlib was designed to provide the ability to process small-footprint full-waveform LiDAR data, provide output products and support the analysis and interpretation of large datasets. The SPD format supports the storage of waveform and discrete return data whilst providing spatial (grid based) indexing of the data for efficient data processing. This system uses pulse waveforms rather than the traditional points as the main data type. This structure allows the full waveform to be processed and stored using the same model in which the data were acquired. It should be noted that at the time of writing SPDlib was still under development. The software is free and distributed under a GPL3 license.

SPDlib is implemented within C++ and was run on a system running Ubuntu (release 11.10) (32-bit). Additional functionality was brought through python bindings. This suite of programs is available to convert LiDAR dataset types, decompose points from waveform data, classify ground returns, interpolate raster height and elevation products and for the calculation of metrics (mostly for vegetation) (Bunting et al., 2011b). Bunting et al. (2011a) outlines the recommended workflow for processing data through the software. Additional steps were required for converting the LAS (1.3) files to the Unsorted Pulse Data (UPD) format, see Figure 4.12. The file conversion was accomplished through using python code which was developed by collaboration between Emma Carolan (Plymouth marine laboratories) and Peter Bunting (Aberystwyth University). The coordinate system was then set to the British National Grid by entering the correct string, as provided by the Geospatial Data Abstraction Library (GDAL) (<http://www.gdal.org/>), using the HDF5 viewer software (The HDF group) (<http://www.hdfgroup.org>).

4.4.1 Full-waveform Gaussian fitting

The SPDlib software utilises Gaussian decomposition (Wagner et al., 2006) to retrieve individual returns. The received power of the waveform (P_i) is linked to the waveforms within the SPD file for each pulse by time (t_i). The zero-crossings of the waveform first derivative above a nominal noise threshold are identified and used as the starting values for N Gaussian amplitude (Q_i) and time (t_i) parameters (Equation 4.10).

$$P_r(t) = \sum_{i=1}^N Q_i e^{-\frac{(t-t_i)^2}{2S_i^2}} \quad [4.10]$$

A baseline threshold needed to be fitted for every entered waveform. Upon inspection of a sample of individually extracted waveforms a single value of 15 was found to account for much of the noise within each waveform. The baseline value of 15 was applied to all waveforms universally. Boundaries were also placed on the pulse width parameter (S_p). An example waveform can be visualised in Figure 4.13. A combination of angular measurements, bearing, positional information of the aircraft and first peak coordinates, trigonometry and the relevant pulse timings (2ns or 1ns) allowed the estimation of the 3D locations for each of the extracted Gaussian peaks fitted.

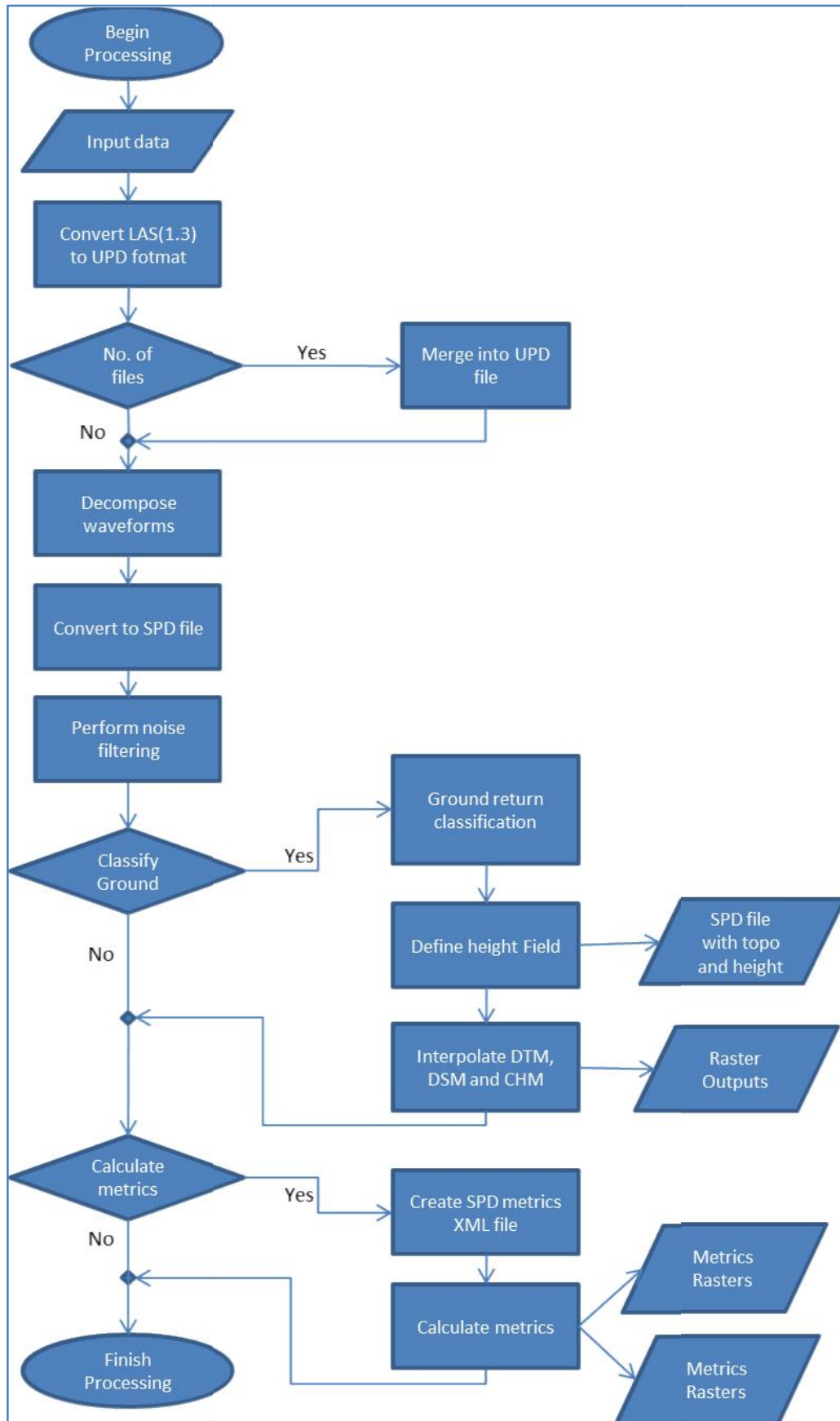


Figure 4.12 – SPDLib airborne waveform LiDAR processing chain. Modified from that presented in Bunting et al. (2011a).

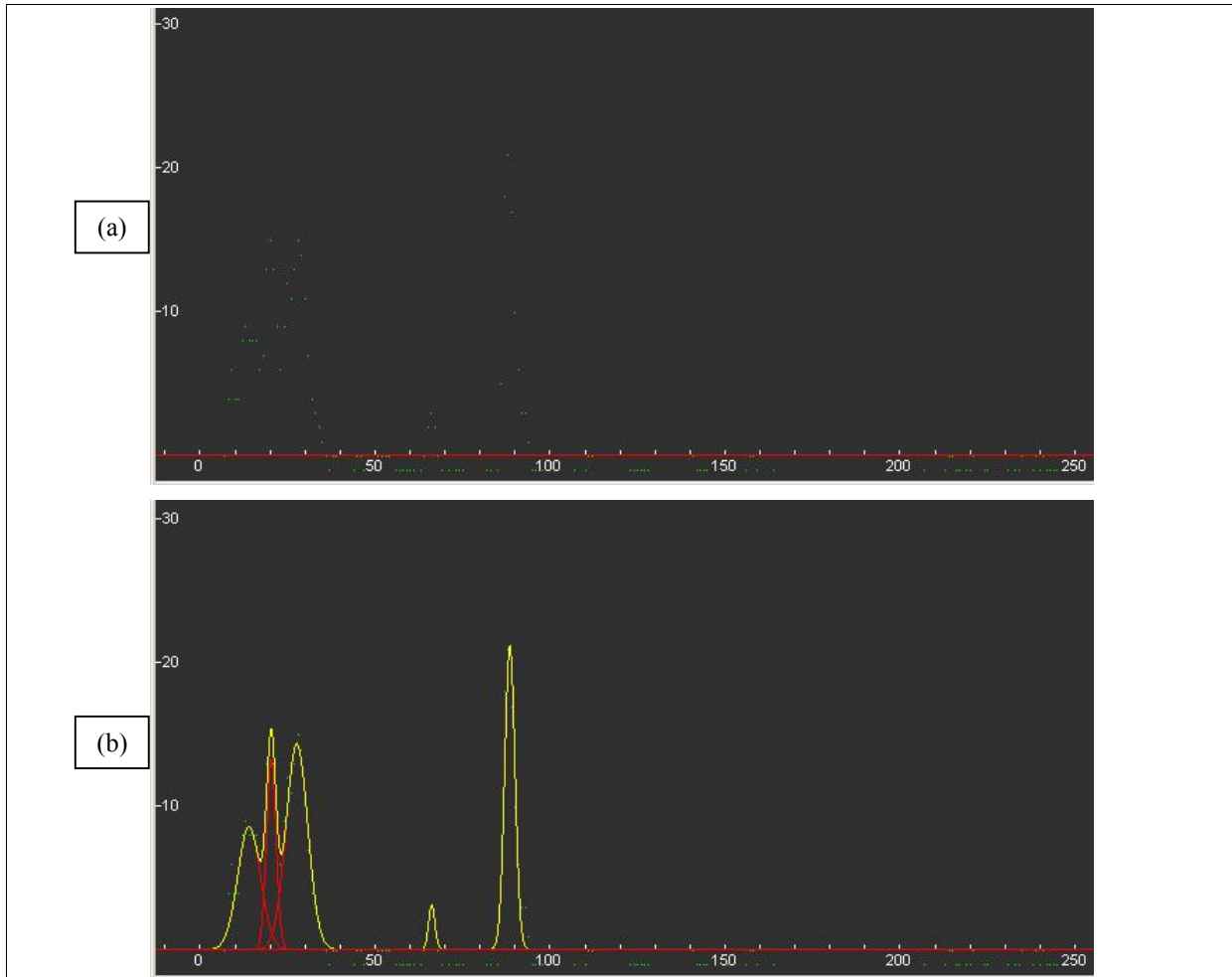


Figure 4.13 – An example of one of the extracted waveforms (viewed in FitYK software). Each of the green dots in the image represents one of the 256 waveform samples (a). Gaussian peaks can be fitted to these data (b) which correspond to the laser pulse interacting with a surface.

4.4.2 Filtering and area-based metric extraction

The classification of those returns which correspond with the ground surface was the next step. The classification algorithm applied was the Progressive Morphological Filtering (PMF) approach by Zhang et al. (2003), which can be applied to landscapes with variable terrain. However as seen in Figure 4.14, it was necessary to include an additional noise filtering processing step to remove some of the vertical errors.

Once the ground returns have been classified the height of each return can be defined in relation to the ground and normalised. The ground is a continuous surface and as such classified points will need to be interpolated into a surface. The natural neighbour algorithm was used for interpolations. As Bater and Coops (2009) demonstrated, this algorithm produced the closest surface to the actual when deriving height surfaces from LiDAR data.

The non-ground heights are then subtracted from the ground heights to remove the effects of the terrain.

Raster outputs of the DTM, DSM and CHM were produced, again interpolated by the natural neighbour algorithm, which were outputted in 32-bit floating ENVI image files. Extracting LiDAR metrics with SPDlib is handled through interrogating an Extensible Markup Language (XML) file. A summary of all extracted metrics is given in Table 4.13. Appendix C sections C.6, C.7 and C.8 contain a full list of extracted metrics. These metrics were output for statistical analysis through applying shapefile inputs of required areas; the table within the shapefile is then populated with the metrics. These metrics were extracted for both leaf-on and leaf-off datasets and the outputs merged, giving a total of 252 metrics. Alternatively, outputs can be set to standard ENVI image files. SPDlib provided the functionality of producing more metrics than RSC LAsTools, such as the number of canopy layers, canopy depth, canopy maximum gap and dominant height.

A greater number of metrics can be extracted from waveform LiDAR data, mainly those relating to the amplitude and width of each of the echoes (or returns) within each waveform. A more detailed description of the information in the surrounding literature concerning these metrics is summarised in Appendix C section C.9. The amplitude is defined as the measurement which characterises the peak power of the echo for each individual return (Wagner et al., 2008), representing the reflection of the laser pulse back in the direction of the receiver. Alexander et al. (2010) state that the amplitude values for a given object would vary depending upon flying height or elevation differences from the emitted laser pulse, even within a single dataset. Thus, as with the DR LiDAR intensity, only metrics related to the distribution of values were utilised, for example skewness and kurtosis.

As mentioned previously backscatter waveforms are popularly modelled by a mathematical function such as the Gaussian function. The extracted parameter representing echo-width can be defined as the width of the fitted Gaussian curve. Again a value of echo-width is generated for each individual return. According to Lin and Mills (2010) the echo-width is related to the slope and the surface roughness of the target that the laser pulse interacts with. More information can be found in Wagner et al. (2006) concerning the definitions of echo-amplitude and echo-width.

The dominant height metric is average height of the non-ground returns in the highest 20% of returns. The metrics of: (i) number of canopy layers, (ii) canopy depth, and (iii) canopy maximum gap were produced through the SPDprofile tool and output in a raster format with 30x30m cell size. The metrics were derived through calculating a two-dimensional vertical profile of returned LiDAR height points per 30x30m grid-cell. This can be conceptualised as height versus number of LiDAR returns. Similarly to the Gaussian decomposition of pulse waveforms, functions were fitted to the vertical profile for each 30x30m cell extent to estimate each of the three metrics. Briefly, the first metric (i) number of canopy layers was determined through calculating the distinct number of groupings within the vertical profile. The second (ii) canopy depth could be calculated by determining the maximum peak width of the peak in the upper stratum. Finally the third (iii) canopy maximum vertical gap, is simply the largest gap, or height range, detected between canopy layers.

It should be noted that the dataset produced from merging the flightlines together caused a number of calculation anomalies to be present in the statistical results. The areas of overlap between lines exhibited drastically different values than for areas where there was no overlap for certain metrics, as illustrated in Figure 4.15. The metrics effected by this related to amplitude and width calculations, such as mean, variance and percentiles. In order to solve this problem the individual flightlines were processed in isolation and the metrics extracted as documented above. The metric shapefiles were then merged together in ArcMap 10. Grid cells in overlapping areas were filtered, removing those grid cells furthest from nadir. No further work was implemented in order to understand this phenomenon, however in the future its understanding should be considered critical.

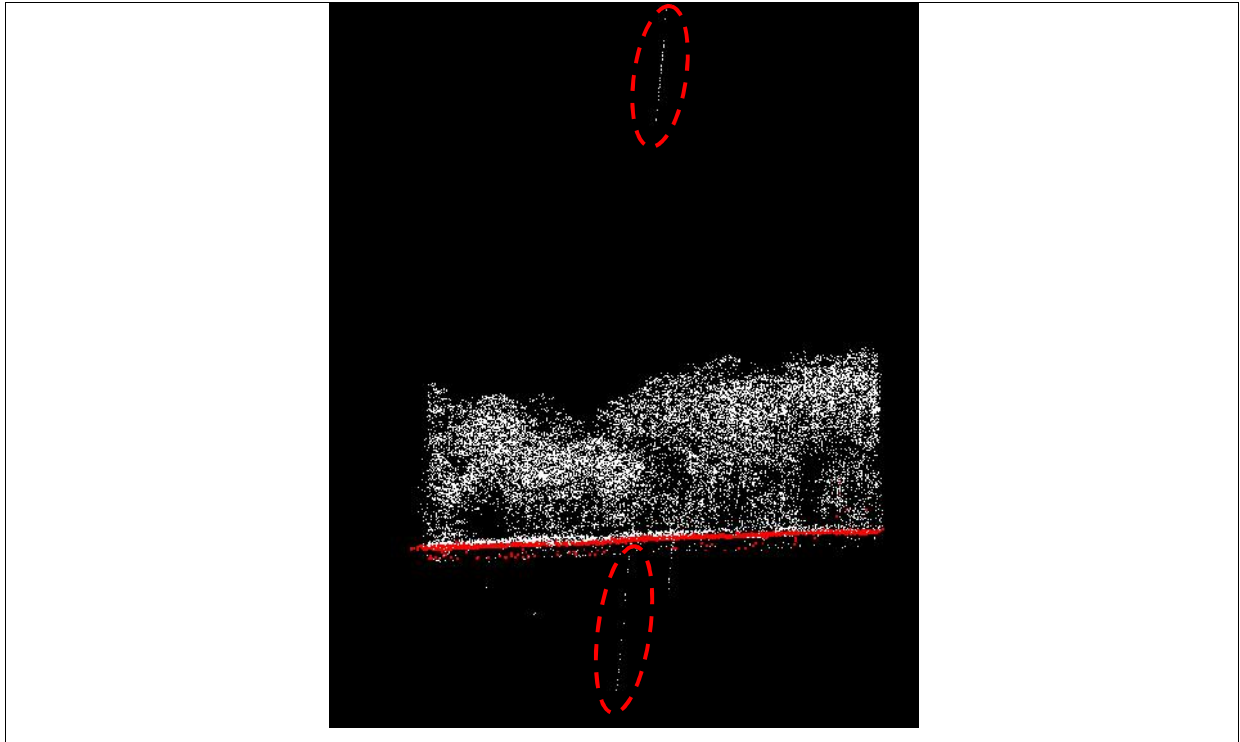


Figure 4.14 – Illustrates the classification of the waveform-derived point cloud with ground (red) and non-ground (white) and a number of errors present in the processing of waveform data.

Table 4.13 – Summary of extracted metrics from waveform LiDAR

Height:	Amplitude:	Width:
Canopy cover	-	-
No. canopy layers	-	-
Canopy depth	-	-
Canopy max. gap	-	-
Dominant height	-	-
-	-	Minimum
Maximum	-	Maximum
Mean	Mean	Mean
Median	Median	Median
Standard deviation	Standard deviation	Standard deviation
Variance	Variance	Variance
Absolute deviation	Absolute deviation	Absolute deviation
Skewness	Skewness	Skewness
Kurtosis	Kurtosis	Kurtosis
Percentiles at 5% intervals (5,10...90,95)	Percentiles at 5% intervals (5,10...90,95)	Percentiles at 5% intervals (5,10...90,95)

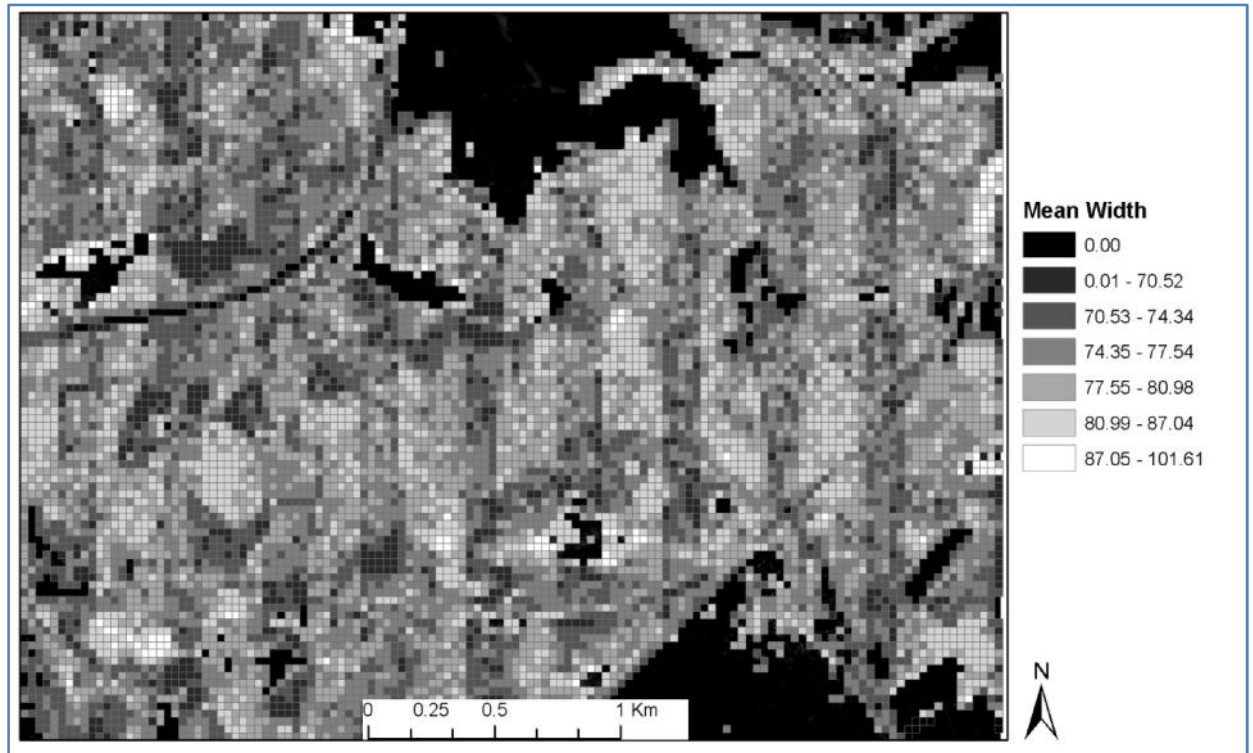


Figure 4.15 – Errors observed in initial FW metric extraction. North to south artefacts are visible due to erroneous values encountered in areas of flightline overlap.

4.4.3 Individual tree detection

In order to compute the same ITC metrics for FW LiDAR as for DR, the SPD format files were converted to LAS 1.2 format (via the ‘spd2spd’ command in the SPDlib software) with the additional amplitude and width metrics removed. They were then input into TIFFS and processed, as in section 4.3.4, to delineate individual tree crowns, and extract metrics related to tree locations, heights and crown dimensions. All 30x30m area summaries were computed as before for the field-plot extents and for the whole study site.

4.4.4 Additional metrics for DR and FW data derived using the FW processing chains

The SPDlib software could also process and extract metrics from the DR LiDAR datasets, utilising much of the same approach identified in the previous sections apart from the waveform decomposition. The SPDlib software could produce a number of additional metrics which could not be calculated through RSC LASTools, namely canopy depth (m), number of canopy layers and the maximum gap between canopy layers (m).

The additional SPDlib metrics (canopy depth, number of canopy layers and the maximum gap between canopy layers) were computed for both the leaf-on and leaf-off DR LiDAR datasets respectively, for both the field plot polygons and the 30x30 m grid-cell polygons. The SPDprofiles and SPDmetrics functions were used for the computation.

4.5Data analysis methods

This section describes in detail the statistical methods implemented in order to, firstly determine the relationships between field plot-level variables to one another and then to derive field-level metrics (as described in section 4.2) from the metrics extracted from the airborne remote sensing datasets. Initially the analysis was conducted on the hyperspectral, DR LiDAR and FW LiDAR data metrics in isolation to one another. The hyperspectral metrics were then combined separately with each of the DR and FW LiDAR data metrics.

It should be noted that a number of metrics recorded in the field relating to animal damage and disturbance, LAI, soil moisture and pH, in addition to the Evenness index were not included in the analysis. These metrics exhibited small ranges between field plot sites and proved to be difficult to model from remote sensing data inputs and map. None of these metrics are required for the selected condition indices.

4.5.1 Field data analysis

The goal of the following procedures was twofold: the first was to assess the similarity of the two field campaign datasets, and secondly was to identify if there were any significant statistical relationships between individual field metrics. Relationships between metrics were investigated using correlation and regression modelling. This process was necessary for assessing the ability to reduce the number of measurements required by field data collection in future works.

4.5.1.1 Comparison of the 2010 and 2012 field populations

In order to be able use the validation field dataset, it was first necessary to determine the overall similarity or dissimilarity between the individual recorded field metrics recorded in 2010 and 2012. This was necessary in order to assess the dataset best used for validation purposes. For the purposes of this research each of the recorded field metrics, such as mean DBH, is considered a population: i.e. population one – the 21 mean DBH values from plots enumerated in 2010; and population two – the 20 mean DBH values from plots enumerated in

2012. Two non-parametric tests, both utilising two independent samples, were utilised to test the two populations, the Mann-Whitney-Wilcoxon and Kolmogorov-Smirnov test. These tests can be used to test if the population distributions are identical without assuming them to follow the normal distribution. Both the Mann-Whitney-Wilcoxon and Kolmogorov-Smirnov tests were implemented through the R software.

The Mann-Whitney-Wilcoxon test is used in experiments in which there are two conditions and different subjects have been used in each condition (Field, 2013), in this case data from two field campaigns. Operating under the assumption the observations are independent of one another, the observations from both groups are combined and ranked, with the average rank assigned in the case of ties. The number of ties should be small relative to the total number of observations. If the populations are identical in location, the ranks should be randomly mixed between the two samples.

The Kolmogorov-Smirnov test is a more general test that detects differences in both the locations and shapes of the distributions between the two populations (Field, 2013). The Kolmogorov-Smirnov test is based on the maximum absolute difference between the observed cumulative distribution functions for both samples. When this difference is significantly large, the two distributions are considered different.

For both of these statistical techniques the null hypothesis was that the two populations were identical. The hypothesis was tested by applying the two independence tests within the R software. The $p < 0.05$ significance level was used, thus if the p value met this criterion the null hypothesis is rejected.

4.5.1.2 Spearman's rho bivariate correlation

In order to determine if there were any correlations between collected field metrics a bivariate correlation analysis was instituted within the SPSS (version 19) (IBM) statistical software. The bivariate correlation method was that of Spearman's rho. The Spearman correlation coefficient is a non-parametric measure of the strength and direction of association that exists between two variables measured on an ordinal scale. Spearman's tests works by first ranking the data, and then applying Pearson's equation to those ranks (Field, 2013). As the direction of the relationship is unknown between variable pairs, all correlations use the two-tailed method.

A correlation matrix was created where every field metric was tested against every other field metric, thus a correlation coefficient and an estimate of significance (p) was calculated for every relationship.

4.5.1.3 Ordinary Least Squares (OLS) regression analysis

Linear regression was used to determine the nature of the relationships between all field plot-level metrics. More specifically this method uses the least squares regression approach. ‘Ordinary Least Squares’ (OLS) means that the overall solution minimizes the sum of the squares of the errors. OLS is a method for estimating the unknown parameters in a linear regression model. A more detailed outline of this approach is available in Field (2013). The resulting relationship can be expressed by a simple formula (see Equation 4.11).

The multiple linear regression approach was implemented through the SPSS software. Multiple forward stepwise regression approach was used, assessing the outputs for evidence of supporting the assumptions of:

- (i) linearity;
- (ii) normality;
- (iii) homogeneity of variance;
- (iv) independence;
- (v) model specification.

Those factors are summarised in Table 4.14.

Table 4.14 – Testing the assumptions of a regression analysis (Chen et al., 2003)

	Assumption/concern:	Description:
i	Linearity	The relationships between the predictors and the outcome variable should be linear.
ii	Normality	The errors should be normally distributed – technically, normality is necessary only for the t-tests to be valid, estimation of the coefficients only requires that the errors be identically and independently distributed
iii	Homogeneity of variance (homoscedasticity)	The error variance should be constant
iv	Independence	The errors associated with one observation are not correlated with the errors of any other observation
v	Model Specification	The model should be properly specified (including all relevant variables, and excluding irrelevant variables)
vi	Influence	Individual observations that exert undue influence on the coefficients
vii	Multicollinearity	Predictors that are highly collinear, i.e. linearly related, can cause problems in estimating the regression coefficients.

Considering the important issue of multicollinearity within this situation, when there is a perfect linear relationship among the predictors, the estimates for a regression model cannot be uniquely computed (Field, 2013). As the degree of multicollinearity increases, the regression model estimates of the coefficients become unstable and the standard errors for the coefficients can get wildly inflated. Through the examination of diagnostic statistics available in SPSS: (i) ‘tolerance’; (ii) ‘variance inflation factor’ (VIF); and (iii) the ‘condition index’, multicollinearity can be detected. Explanations of these metrics can be found in Field (2013).

Each of the various diagnostics was used to test the correlations between all of the field-plot metrics. If an assumption was found to be false, of concern or to be out of tolerance, the regression was re-run and one or more predictor values removed in order to improve the regression. The output was then assessed again.

The regression result was then assessed for its significance using an Analysis of Variance (ANOVA) F-test and the student’s t-test. Here if a value of p is below 0.05, it is reported as significant for both F- and t-values (Chen et al., 2003; Field, 2013). If a coefficient was not significant, it was dropped from the regression. Additionally, the Residual Standard Error (RSE) was checked for each model; out of all the models produced those which minimised the RSE were selected.

In order to get an indication of how much of the variance encountered in the dependent variable was accounted for by the regression model, the R-squared statistic was utilised. This is an overall measure of the strength of association and does not reflect the extent to which any particular independent variable is associated with the dependent variable.

The output from the regression model yields the inputs for the regression equation. These are the values in the regression equation for predicting the dependent variable (Y) from the independent variables (x). The regression equation is presented as:

$$Y_{\text{predicted}} = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n \quad [4.11]$$

where b_0 is the intercept, $b_1 \dots b_n$ is the coefficient which corresponds to the independent (predictor) variable $x_1 \dots x_n$ (e.g. b_1 corresponds to x_1). This equation was generated for each individual field plot attribute, as predicted by combinations of other field plot metrics using multiple forward stepwise regression. Non-significant metrics were removed from the

analysis to produce minimum adequate models. Additionally, efforts were made to limit the standard errors of the best models.

4.5.2 Hyperspectral analysis

As outlined in section 4.4, metrics extracted from hyperspectral datasets totalled 5 from the ITC classification and 149 from the spectral indices (154 in total). This combined both leaf-on and leaf-off datasets. Below is a description of various statistical approaches used to derive a statistical method of predicting field-level attributes from hyperspectral data. Hyperspectral statistics were derived for the 30x30m areas corresponding to the field-plot locations. Data from the 21 field-plots visited in 2010, one attribute at a time, were then regressed against the hyperspectral values as a means of generating the required equations for predicting the field plot-level metrics over the whole study site.

4.5.2.1 OLS linear multiple regression

As in section 4.5.1 an OLS regression analysis was performed in order to determine the nature of the relationships between a field plot-level metric (dependent variable) and one or more hyperspectral-derived metrics (explanatory variables) using SPSS. Multiple forward stepwise regression was used as described above, to produce minimum adequate models predicting field plot-level attributes from the hyperspectral-derived metrics.

4.5.2.2 Akaike's Information Criterion

As an alternative method of deriving the regression relationships the Akaike's Information Criterion (AIC) technique was explored. AIC in essence balances the number of parameters and fit to the data (likelihood). This technique was implemented using the R statistical software. AIC is a measure of the relative 'goodness of fit' of a statistical model and is defined by the equation of:

$$\text{the relative 'goodness of fit'} \quad [4.12]$$

$$AIC = 2k - 2\ln(L)$$

where k is the number of parameters in the regression model, and L is the maximized value of the likelihood function for the estimated model. A small value of AIC indicates a better combination of simplicity and fit to the data.

Given the relatively small sample size ($n=21$) and large predictor size ($k=196$) it was necessary to consider a modification to 'AIC', termed 'AICc' (Burnham and Anderson, 2002). AICc is defined as:

$$AICc = AIC + \frac{2k * (k + 1)}{n - k - 1} \quad [4.13]$$

where n denotes the sample size. Therefore, AICc includes a penalty correction for extra parameters.

Within the R software the “MuMin” (Multi-Model Inference) package (Barton, 2012) was used to run the AICc analysis, using the ‘Dredge’ function. This function generates a set of models with combinations (subsets) of the terms in the global model, with optional rules for inclusion. The function runs through each possible combination of variables in order to derive the most significant regression equation which accounted for the most variance. Unfortunately the statistical tool could not accept more than 30 input metrics due to computer memory limitations.

The number of possible combinations of predictor variables (≥ 155) presented a number of problems as the number of potential permutations was vast. There was therefore a high risk of identifying spurious relationships. Below is an adaption of the methods outlined in Langton et al. (2010) and Burnham and Anderson (2002). To avoid this problem a further phase of modelling was carried out in an attempt to identify those variables which would be significant while reducing the potential for collinearity.

A ‘data mining’ exercise was conducted in order to investigate other important predictor variables. To determine which variables had the most potential for prediction of forest attributes, automatic stepwise AICc selection was used on a subset of six random predictor variables for 500,000 iterations. If a predictor variable was significant, it was recorded for each of the iterations. Each application of an AICc model was assessed using ANOVA test. Each input variable had a corresponding F-test and p -value. As before, a variable was considered significant if $p \leq 0.05$. The result of this process was a table summarising which variables were significant for each of the random subset selection iterations. A results table was then produced where a count for each time an attribute was significant was calculated as a measure of which attributes were of most relevance to a given field plot-level variable. A full list of this R code for this task is presented in Appendix D section D.1. This was applied for each of the field-level variables.

Following this process, the 20 predictor variables with the highest counts were input into a further AICc process in order to derive a regression equation. If variables were known *a priori* to have no relation to the independent variable they were removed and the next best predictor variable was added. At this stage a limit was imposed on the number of predictor variables allowed into the stepwise AICc regression in each step. A maximum of 6 of the 20 predictor variables could be entered in any single iteration of the model, this was in order to account for processing time and system memory limitations. The AICc was run adding in each possible combination of 1-6 variables. The delta-AIC value of each model was then assessed in order to determine the likelihood of the candidate model. When delta-AIC was less than or equal to 2, the given model was suggested to be within the range of plausible models that best fit the observed data (Burnham and Anderson, 2002). Therefore any model with an AIC above 2 was discounted.

As in the previous section, a number of diagnostic tests were applied to assess the regression assumptions in this model, as listed in Table 4.14. Significance tests such as the ANOVA and students t-test were available as before. Diagnostics designed to detect multicollinearity were VIF and tolerance. These were calculated using the ‘Faraway’ package in R (Faraway, 2011). The condition index was calculated using the ‘perturb’ package (Hendrickx, 2012). A full listing of this R code is presented in Appendix D section D.2. If the model failed these diagnostic tests, the variable(s) identified as not significant and/or collinear were removed from the analysis and the AICc procedure re-run. Efforts were made to limit the standard error of the model. AICc delta and weight values were reported for each model.

4.5.3 Discrete-return and full-waveform LiDAR analysis

As outlined in section 4.4 a total of 196 and 261 attributes for DR and FW data respectively, were extracted for the combined leaf-on and leaf-off datasets, including those derived from ITC centroids. These attributes included those metrics which could be directly related to fieldplot-level metrics and those which required statistics to determine a relationship.

Below is a description of various approaches used to derive a statistical means of predicting field-level attributes from LiDAR data. LiDAR statistics were derived for the 30x30m areas corresponding to the field-plot locations. Data from the 21 field-plot sites recorded in 2010 were regressed against the LiDAR values one attribute at a time to generate equations for predicting the field plot-level metrics over the whole study site.

4.5.3.1 OLS linear multiple regression

An OLS regression analysis was performed to determine relationships between a field plot-level metric and one or more LiDAR derived metrics, using SPSS. A multiple forward stepwise regression approach was used, assessing the outputs for evidence of supporting the assumptions summarised in Table 4.14. It should be noted that multicollinearity between LiDAR attributes was of critical concern. As before, non-significant metrics were removed from the analysis and the regression process re-run and efforts were made to limit the standard errors of the models.

4.5.3.2 Akaike's Information Criterion

As in section 4.5.2.2, a two stage AICc procedure was implemented, in order to first identify those metrics most likely to be correlated with a field-level metric. The second step was to perform an AICc stepwise regression approach to model the DR and FW LiDAR metrics against field-level metrics. To reiterate, the field data collected in 2010 were used, in addition to the remote sensing metrics extracted for the same spatial extents for both leaf-on and leaf-off data. The assumptions summarised in Table 4.14 were tested. Students-t test, ANOVA, VIF and condition indices were utilised. Efforts were made to limit the standard error of the model. AICc delta and weight values were reported for each model.

This procedure was applied to the DR and the FW datasets separately to produce predictive equations relating the LiDAR variables to the field plot-level attributes.

4.5.3.3 Direct calculation of indexes from individual tree detection data

Two of the condition assessment indices, as listed in section 4.1.3, require the explicit measurement of individual trees in order to be calculated, these were the Vertical Evenness (VE) index (Neumann and Starlinger, 2001) and the Clark-Evans Aggregation index (Clark and Evans, 1954). The R software was used to calculate statistics based on the individual tree crown (ITC) centroids as derived through TIFFS. Each ITC point contained coordinates and an estimate of height. Custom R script was developed to estimate the VE index by stratifying the individual heights of the ITC objects based on the maximum height recorded in the 30x30m plot and applying a modified Shannon Index. This R script is documented in Appendix C section C.10. The Aggregation index was calculated using the distances between ITC point coordinates and an estimate of the number of tree stems within the plot. The code is documented in Appendix C section C.11.

4.5.4 Combined hyperspectral and LiDAR analysis

Hyperspectral metrics were combined first with DR LiDAR and then with FW LiDAR in isolation from one another. This combined both leaf-on and leaf-off datasets. Due to the large number of potential predictor variables in the combined datasets, there was a concern that this could cause problems with both multicollinearity and the limitations of the available computer hardware.

4.5.4.1 Defining combined variable subsets

Hyperspectral-derived data and DR LiDAR metrics, of 155 plus 196 variables, respectively (total 351), were processed in the aforementioned manner. Likewise, hyperspectral derived data and FW LiDAR metrics, of 155 plus 261 variables, respectively (total 416) were processed by the same approach.

To reduce the number of variables whilst retaining key predictors, for the combination of remote sensing variables a modification of the method outlined in section 4.5.2.2 was adopted. To reiterate, the first part of the procedure was to perform a ‘data mining’ exercise using 500,000 iterations to identify the variables likely to be the most significant. This procedure was used for derived hyperspectral, DR and FW LiDAR datasets. To combine the two remote sensing datasets the 20 most frequently selected remote sensing variables for each field metric were extracted for each of the two combined datasets.

4.5.4.2 OLS linear and AICc regression

OLS regression analysis and AICc regression were then performed as described above in sections 4.5.1 and 4.5.2.2. For the OLS analysis, a total of 20 metrics from hyperspectral derived data and 20 from DR/FW LiDAR were identified for each of the field-plot metrics. For the AICc, the process was modified to include only 10 metrics from each of the two remote sensing sources, i.e. 10 from the hyperspectral and 10 from the LiDAR datasets. Exceptions were made to this rule to include those remote sensing metrics identified as significant in the previous AICc selection procedures conducted for both hyperspectral and LiDAR analysis in isolation (see sections 4.5.2.2. and 4.5.3.2). The number of values input into the AICc procedure was limited to a maximum of 24 due to computer memory limitations. The R-script developed to perform this task is documented in Appendix D section D.3.

4.6. Validation using field data

The 2012 field data was acquired for the explicit reason of validation of the predictions made through the various statistical methods reported in the previous (sections 4.2 to 4.5).

4.6.1 Using remote sensing to predict field plot-level metrics

The coordinates of the four corners of the 2012 field plot locations were used to produce Esri format shapefiles, these extents were then used to extract the required remote sensing metrics from the three datasets. Direct comparison of field metrics and remote sensing derived field plot-level metrics could be made for a small number of metrics, summarised in Table 4.15.

Table 4.15 – Field measurements and remote sensing metrics for comparison

No.	Field Metric:	RS Metric:
1	Canopy height (mean, max.) (m)	Mean TIFFS height (m) (LiDAR); Dominant Canopy Height (m) (LiDAR).
2	Canopy layers	Leaf-on and leaf-off canopy layers (LiDAR).
3	Overstorey species present (height >15m)	Species classifications from ITC count (Hyperspectral/LiDAR).
4	Number of overstorey trees (height >15m)	ITC count (Hyperspectral/LiDAR)

The remote sensing metrics were then extracted by the methods outlined in the previous sections for the extents of the 2012 field work polygons for hyperspectral, DR and FW LiDAR datasets. These metrics were used as inputs to the various statistical equations generated in section 4.5 for the prediction of field plot-level metrics. Comparisons between validation 2012 field work data and airborne remote sensing derived plot-level metrics were made for each of the field measured plot-level metrics outlined in section 4.1.

An R-script was developed in order predict the field plot-level metric from remote sensing data from each of the 10 statistical models (5 datasets and 2 regression approaches); this is documented in Appendix E section E.1. Once the relevant remote sensing metrics were extracted for the spatial extents of the 2012 validation fieldwork polygons, the regression model equations were applied in order to predict field plot-level information.

A comparison of the field plot metrics with the estimated values from the remote sensing derived sources was then performed. The Root Mean Square Error (RMSE) and the Normalized Root Mean Square Error (NRMSE) were computed for each of these comparisons as a measure of accuracy. These functions are defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}} \quad [4.14]$$

where x_i is the predicted value and y_i is the value recorded by fieldwork. The product of subtracting y_i from x_i is known as a residual. Squaring the residuals for the population, averaging the squares, and taking the square root gives the RMSE.

Normalized Root Mean Square Error (NRMSE) was calculated in order to provide a scale-independent measure of accuracy. This is defined as:

$$NRMSE = \frac{RMSE}{y_{max} - y_{min}} \quad [4.15]$$

RMSE is divided by the range of observed values, where field observed maximum is y_{max} , and observed minimum is y_{min} .

4.6.2 Computation and validation of condition indices

Condition indices were computed from the remote sensing derived field-level metrics and compared with those calculated from the field data. The remote sensing derived field metrics produced as part of the previous step (section 4.5) with the lowest relative RMSE and NRMSE were selected as inputs into the various condition index methods. The process was carried out using the best models from any of the five datasets, as defined by the lowest RMSE/NRMSE. The construction of the indices is described in section 4.1.3.

The eight indices tested are as follows: (i) Shannon index; (ii) Simpson index; (iii) Clark-Evans aggregation index (Clark and Evans, 1954); (iv) tree diameter differentiation index (Füldner, 1995, cited in Vorčák et al., 2006); (v) vertical evenness (VE) index (Neumann and Starlinger, 2001); (vi) the complexity index (Holdridge, 1967); (vii) complex stand diversity index (CSDI) (Jaehne and Dohrenbusch, 1997, cited in Newmann and Starlinger, 2001, Vorčák et al., 2006); and (viii) a scoring method, as defined by Cantarello and Newton (2008). The construction of these indices is documented in section 4.1.3.

The Shannon and Simpson indices were calculated directly from hyperspectral data in addition to using statistical means. It should be noted that a number of indices did not use regression derived inputs, and will not be discussed here. For example, the VE and aggregation indices were computed from ITC information extracted from DR LiDAR data due to the requirements for the condition index to use individual tree data.

Tree diameter differentiation and the CSDI required additional statistical models to be created for their inputs. The former required the computation of the sum of the differences in DBH of stems within the field plot. The latter required the sum of the three largest and three smallest DBH values recorded within the field plot to be calculated. Both of these metrics were estimated through the regression approaches defined previously for all five datasets. The model estimates with the smallest RMSE/NRMSE were selected for input into the condition index calculation.

The CSDI required a number of metrics derived from regressions and extracted directly from both the hyperspectral and LiDAR ITC objects in order to be computed for three of four of its component indices. The first of these related to the estimation of the relative proportions of the tree species with the highest population to that of the lowest population. This was calculated using the classified ITC objects produced in eCognition, where each ITC object was considered as an individual tree stem. The second component index, the index of vertical separation, was estimated statistically using regression, as defined before. Component index 3 required estimates of the three minimum and three maximum nearest neighbour distances in-between DR LiDAR ITC objects. Component index 4 required the two minimum and two maximum tree crown diameters from DR LiDAR ITC data, in addition to the ‘best’ regression estimates (i.e. with the lowest RMSE), for estimates of the stem count, height to the lowest live branch and standard deviation of the height to the first live branch.

Modifications were made to two of the components of the complex stand diversity index: (CSDI-3) the index of spatial distribution and (CSDI-4) the index of crown differentiation. These modifications were necessary to take account of some of the initial findings relating to the LiDAR TIFFs ITC outputs and statistical estimates of forest metrics.

For CSDI-3 these were the underestimation of detecting individual trees, and the overestimation of the distance between nearest neighbouring (NN) trees using TIFFS ITC data. The distances between NN trees as derived from ITC centroids were modified by subtracting two times the standard deviation of NN distances (derived from statistical outputs, not ITC metrics) from the actual ITC distances. This was done so the values better resembled 2010 fieldwork measurements for NN distances.

The index for crown differentiation required an estimate of the lowest possible height of the first live tree branch (a.k.a. crown base height); a simple surrogate was constructed subtracting the standard deviation of the height of the first live tree branch from the plot average height to the first live branch. If the height was below an arbitrary 2m threshold the average height to the first live branch was used. The threshold was instituted to reduce the possibility of a negative height value being entered into the calculation.

An R script was developed to calculate all the condition indices:

- (i) the tree diameter differentiation index;
- (ii) the complexity index;
- (iii) the CSDI; and
- (iv) the scoring index (Cantarello and Newton, 2008).

The R-script developed for this purpose is documented in Appendix E section E.2. The fieldwork derived index values were then compared with the remote sensing derived values, where RMSE and NRMSE were calculated.

4.7 Mapping the results across the study site

Utilising the 30x30m grid polygon produced using the GME software each of the distinct remote sensing metrics could be mapped. There were approximately 32,100 30x30m grid-cells within the study site extent. Hyperspectral index data were added to the grid shapefile along with overstorey species number and TIFFS ITC summary metrics relating to stem number, average height, crown area and crown volume within ArcMap. DR LiDAR metrics were obtained through the RSC LAStools software using a 30x30m pixel sized raster grid for each of the extracted metrics aligned to the coordinates used in the polygon cells. They were extracted using a combination of generating a central point within the polygon cell, sampling the pixel value which intersected with this point and adding the value to the polygon grid via a spatial join. SPDlib metrics for FW LiDAR were output directly into the polygon shapefile.

Each of the remote sensing attributes were then mapped at the 30x30m resolution for the whole of the study site, these datasets are summarised in Table 4.16. Once the field-level attributes were predicted by the various statistical methods, estimates of field-level information could be made for each 30x30m shapefile extent through the application of an appropriate regression model formula. A total of ten potential model formulae existed for each of the field-level metrics to be estimated.

Table 4.16 – Summary of inputs for full study site

Datasets foreach 30x30m area:		
Hyperspectral	DR LiDAR	FW LiDAR
<u>eCognition (and R) output:</u> <ul style="list-style-type: none"> Population composition summary; Diversity indices from classified over-storey. 	<u>RSC LAStools outputs:</u> <ul style="list-style-type: none"> Height and intensity metrics for ground, vegetation and all returns [for both leaf-on and leaf-off data]. <u>SPDlib outputs:</u> <ul style="list-style-type: none"> Canopy depth; canopy max vertical gap and no. canopy layers. 	<u>SPDlib outputs:</u> <ul style="list-style-type: none"> Height, amplitude and width metrics for ground, vegetation and all returns [for both leaf-on and leaf-off data]. Canopy depth; canopy max vertical gap and no. canopy layers.
<u>Hyperspectral Index area summaries</u> <ul style="list-style-type: none"> Summary of 13 spectral indicates [for both leaf-on and leaf-off data]. 	<u>TIFFS outputs:</u> <ul style="list-style-type: none"> No. of stems; Tree height summary; Crown parameter summary; VE index estimate; Aggregation index estimate. 	<u>TIFFS outputs:</u> <ul style="list-style-type: none"> No. of stems; Tree height summary; Crown parameter summary; VE index estimate; Aggregation index estimate.

In order to apply each of the model equations to the appropriate 30x30m cell, the R-script documented in Appendix E sectionE.1 (also in section 4.6.3) was modified to apply the calculation to all 30x30m grid cells across the study site extent. The output from R was a database (.dbf) file which could be combined with the grid shapefile using the Cell ID attribute. Once completed each of the polygon grid cells could be used to map field attributes, such as mean DBH(m) and standing deadwood volume (m³) across the whole study site.

Finally, it was necessary to calculate the condition indices using the outputs of the previous calculations applied to the remote sensing metrics. In relation to the validation work (section 4.6.3), the most accurate model equations were selected for input into the condition index calculations. An R script was developed to extract the relevant field estimate inputs from the 30x30m polygon grid cells as inputs to condition index calculations. The R script presented in Appendix E sectionE.2was modified to account for the greater number of polygons for which estimates were required. Multiple indices were calculated through the R script, these were: (i) tree diameter differentiation (Füldner, 1995); (ii) the complexity index (Holdridge, 1967); (iii) CSDI (Jaehne and Dohrenbusch, 1997, in Newmann and Starlinger, 2001, Vorčák et al., 2006) and (iv) the scoring method, as defined by Cantarello and Newton (2008). The Shannon, Simpson, Evenness, VE and aggregation indices were calculated using modifications to the R-scripts documented in Appendix sections B.2, C.10 and C.11.

Once the R-script calculations were completed, the condition index values were output from R in the database (.dbf) format and combined with the 30x30m polygon grid in ArcMap by linking the Cell IDs. Thus, each grid cell within the study-site map contained an estimate of each indicator metric and the nine predicted index values.

4.8 Methods summary

The methods chapter began with an overview of the field datasets for the 41 field plots, 21 of which were visited in 2010 and 20 visited in 2012. Various area-based summary metrics were calculated for each field-plot, to produce estimates such as mean DBH, basal area and canopy openness. In addition, nine condition indices relating tree species or structural diversity, or both, were calculated for each field plot. A total of 39 metrics were generated for each field plot.

The hyperspectral datasets required a number of pre-processing steps in to make the data suitable for analysis. These steps involved the application of radiometric, geometric, and atmospheric corrections to the hyperspectral imagery, followed by mosaicing the individual flightlines together. Once pre-processing steps were complete a total of 13 VI were then calculated from the leaf-on and leaf-off datasets. These VIs related to the greenness, light use efficiency, leaf pigments, and canopy water content.

A process to reduce the high-dimensionality of the hyperspectral data was instated, employing concepts such as PCA and MNF. This was to reduce the dataset size and reduce noise and extraneous data. Hyperspectral data processed using MNF was selected as the best approach. This dataset was then classified using an object-based classification method. The object-based method employed a combination of ITC delineation algorithms to identify ITC objects, then employed hierarchical classification using fuzzy class membership rules. Summary metrics were produced for each of the field plot areas for species classification and VI metrics.

The DR LiDAR data required various pre-processing steps to be implemented to filter the LiDAR point cloud to identify ground and non-ground points (using a progressive morphological filter), and then to normalise each point's height information to above-ground height only. Again, the flightlines were mosaiced into one file. Various area-based metrics were extracted from the DR LiDAR dataset. A total of 187 area based statistics were

extracted from leaf-on and leaf-off data for both height distribution and intensity metrics for each of the plot areas. Metric examples include the mean height, canopy closure and the skewness of LiDAR intensity values.

ITC analysis was performed upon the leaf-on 2010 DR LiDAR data, where estimates of tree height, crown radius and canopy geometric volume could be made for each tree. These metrics were averaged or totalled for area-based metrics, for example, an estimate of the number of trees per 30x30m area could be made.

FW LiDAR data presented a number of unique problems in terms of processing it to a level to make is usable for analysis. This dataset type required additional pre-processing steps to derive a ‘traditional’ point cloud. Therefore a method utilising Gaussian decomposition was used to identify individual returns from each of the pulse waveforms, and calculate its position in 3D space. Once complete, filtering of ground and non-ground points could be performed, again using a progressive morphological filter, followed by normalising the point heights to above ground height. The FW dataset allowed additional area-based metrics to be calculated based on the size and shape of the waveform response to the return, i.e. wave amplitude and width, in addition to height information. A total of 252 metrics were extracted for both leaf-on and leaf-off datasets.

Once the metrics had been extracted from the various field and remote sensing datasets a statistical analysis was implemented. Beginning with the field dataset, three statistical techniques were applied. The first was in order to determine similarity or dissimilarity of the data collected in the two field campaigns. The second, to determine if any of the field metrics correlated with any of the others, using a Spearmans rho bivariate correlation analysis. The third technique was to employ OLS multiple regression in order to determine if the individual field metrics were related to any combinations of the other field metrics.

Regression of remote sensing metrics against 39 field data metrics was done to develop predictive models to estimate these field data metrics in other forest areas assessed by the remote sensing data. Both OLS and AICc regression approaches were used to test the five datasets, these were: (i) hyperspectral metrics; (ii) DR LiDAR metrics; (iii) FW LiDAR metrics; (iv) hyperspectral combined with DR LiDAR metrics; and (v) hyperspectral combined with FW LiDAR metrics.

Fieldwork plots recorded in 2012 were used for validating the OLS and AICc regression models for the 39 field plot-level metrics. Thus, each of the remote sensing metrics identified in the statistical models were extracted for the extents in the validation field plots. The regression model equations generated in the previous steps were applied to this dataset. RMSE and NRMSE were calculated for each result of the statistical model estimates. In addition, a number of measurements which could be extracted directly from the data, such as tree counts and canopy openness, were tested.

The same approach was applied to the calculation of condition indices, where field calculated index values were compared with those calculated from indicators derived from regression model estimates using remote sensing metrics. RMSE and NRMSE values were calculated for the 20 validation field plot-extents.

Slight modifications were made to the workflow employed in the methods of validation to apply the regression model calculations to the entire study site, where remote sensing area-based metrics were extracted for a regularly spaced grid of 30x30m areas. A number of mapping products were created for estimates of condition indicator values such as mean DBH or species richness across the whole study site. Condition indices were calculated from the estimated indicator values and validated against 2012 field data. As before the condition index values were then mapped across the study site.

Chapter 5 – Fieldwork results

This chapter reports the results of the various forest structural and compositional metrics collected during fieldwork. It includes:

1. an overview of fieldwork data characteristics;
2. any significant statistical relationships between field recorded metrics;
3. the presentation of calculated condition indices from the field data.

It should be noted that the field data collected in 2010 were used to build the models documented in subsequent chapters, whereas field data collected in 2012 were used to validate these models.

5.1 Overview of fieldwork data

This section discusses the results of both the summer 2010 and 2012 field campaigns, as identified in Chapters 3 and 4. Table 5.1 summarises the fieldwork data for the field plots recorded in 2010, while Table 5.2 summarises the fieldwork data for the field plots recorded in 2012 to illustrate some of the differences per plot. Species richness varied between one and four for over-storey species in the sample plots. Many of the plots visited in the fieldwork campaign had very few instances of regenerating tree species in terms of saplings or seedlings, and therefore the native seedling SH values were low. It is worth considering that these small population sizes may influence the strength of the produced regression models. Between sites, vascular plant species richness was comparable between mixed, coniferous and deciduous plots. Overall however vascular plant diversity was low. The volume of standing and fallen deadwood was greater in deciduous woodland. As expected, tree DBH size varied significantly between trees within the older growth deciduous plots, and was significantly less in the conifer plantation plots. The proportion of larger trees also followed this pattern.

The metrics recorded for each of the plots exhibit a different spatial structure, and a pattern can be seen between mean DBH and height (and to a lesser extent spacing). However, total downed deadwood seems to be unrelated to tree size.

Table 5.1 – Summarised Descriptions for field plots visited in 2010

Plot location	Tree Species Layer	Regeneration & understorey	Deadwood Amount
#1 E 435820; N 102950. (Hawkhill) *Plantation* (Soil pH 5.1)	Species: Scots Pine; Silver Birch; and Beech. Spacing: low (2.5m). Canopy openness 20.5%. Mean height is 17m. Mean DBH is 30cm.	Species: Silver Birch; Holly, Density: low Understorey: ferns (45%) and grass (54%). *Animal browsing.	No standing dead and few fallen trees. Downed deadwood volume is low (10 m ³ ha ⁻¹) in none to moderate decay state.
#2 E 435621; N 102857 (Hawkhill) *Plantation* (Soil pH 5.7)	Species: Corsican Pine; and Oak. Spacing: low (4m). Canopy openness: 17%. Mean height is 18.3m. Mean DBH is 34cm.	Species: Holly Density: low Understorey: ferns (80%), grass (15%), foxgloves (0.5%) and moss (4%). *Animal browsing	No standing or fallen dead trees. Downed deadwood volume is low (13 m ³ ha ⁻¹) in a moderate to high decay state.
#3 E 435794; N 102614 (Hawkhill) *Plantation* (Soil pH 6.0)	Species: Corsican Pine; Birch; and Oak. Spacing: low (4.5m). Canopy openness: 19.6%. Mean height is 18m. Mean DBH is 36cm.	Species: Holly. Density: low. Understorey: grass (80%), Gorse (3%), and fern (15). *Animal browsing.	No standing or fallen dead trees. Downed deadwood volume is low (2 m ³ ha ⁻¹) in a moderate to high decay state.
#4 E 435543; N 103100 (Frame Wood) *Semi-ancient* (Soil pH 6.0)	Species: Oak; Holly; and Silver Birch. Spacing: low (4.8m). Canopy openness: 21%. Mean height is 17.5m, Mean DBH is 57cm.	Species: Holly. Density: low. Understorey: fern (50%), grass (20%), and moss (5%). *Animal browsing.	Standing dead found (47 m ³ ha ⁻¹) in moderate decay state. Downed deadwood volume is low (9 m ³ ha ⁻¹) in a moderate to high decay state.
#5 E 434808; N 101918 (Hawkhill) *Plantation* (Soil pH 5.7)	Species: Douglas Fir; Scots Pine; and Silver Birch. Spacing: low (3.5m). Canopy openness: 23%. Mean height is 24m. Mean DBH is 34cm.	Species: Holly. Density: low. Understorey: grass (90%), and fern (5%).	Standing dead found (0.8 m ³ ha ⁻¹) in a moderate decay state. Downed deadwood volume was low (8.6 m ³ ha ⁻¹) in a moderate to high decay state.
#6 E 434549; N 101919 (Hawkhill) *Semi-ancient* (soil pH 5.5)	Species : Oak; and Holly. Spacing: high (8.5m). Canopy openness: 13%. Mean height is 33m. Mean DBH is 62cm.	Species: Holly. Density: low. Understorey: fern (50%), nettles (5%), with bluebells, moss and nettles.	Standing dead found (47 m ³ ha ⁻¹) in moderate decay state. Downed deadwood volume is low (10 m ³ ha ⁻¹) in a high decay state.
#7 E 435071; N 102558 (Hawkhill) *Plantation* (Soil pH 5.2)	Species: Douglas Fir; Oak; and Silver Birch. Spacing: moderate (5.2m). Canopy openness: 26%. Mean height is 22m. Mean DBH is 40cm.	Species: Holly; Douglas Fir, Silver Birch; and Red Cedar Density: High. Understorey: grass (30%), Fern (20%), with moss (10%) and heather.	No standing dead and few fallen trees. Downed deadwood volume is low (16 m ³ ha ⁻¹) in none to moderate to high decay state.
#8 E 435021; N 102399 (Hawkhill) *Plantation* (Soil pH 5.7)	Species: Douglas Fir; Norway Spruce; and Silver Birch. Spacing: low (2.9m). Canopy openness: 22%. Mean height is 17m. Mean DBH is 33cm.	Species: Holly; and Norway Spruce. Density: low. Understorey: grass (60%), and fern (30%).	Standing dead found (3 m ³ ha ⁻¹) in a moderate decay state. Downed deadwood volume was low (3 m ³ ha ⁻¹) in a moderate decay state.

Table 5.1 (continued)

Plot location	Tree Species Layer	Regeneration & understorey	Deadwood Amount
#9 E 436549; N 103213 (Tantany Wood) *Semi-ancient* (Soil pH 6.1)	Species: Oak; and Beech; Field Maple; hawthorn; and Holly. Spacing: moderate (5.3m). Canopy openness: 15%. Mean height is 28m. Mean DBH is 48cm.	Species: Holly, Beech and Ash. Density: moderate. Understorey: (sparse) grass (3%), and fern (3%). *Animal browsing	High volume of standing dead ($112 \text{ m}^3 \text{ ha}^{-1}$) in a moderate to high decay state. Downed deadwood volume was high ($98 \text{ m}^3 \text{ ha}^{-1}$) in a moderate to high decay state.
#10 E 436351; N 103769 (Tantany Wood) *Semi-ancient* (Soil pH 6.5)	Species: Oak; and Beech; and Ash. Spacing: low (2.6m). Canopy openness: 15%. Mean height is 31m. Mean DBH is 50cm.	Species: Holly; and Ash Density: moderate. Understorey: (sparse) grass (20%), and a mix of sorrel, moss and ground *Animal browsing Ivy.	High volume of standing dead ($77 \text{ m}^3 \text{ ha}^{-1}$) in a moderate to high decay state. Downed deadwood volume was high ($162 \text{ m}^3 \text{ ha}^{-1}$) in a moderate decay state.
#11 E 436793; N 103778 (Tantany Wood) *Semi-ancient* (Soil pH 5.9)	Species: Oak; Silver Birch; Yew; Alder Buckthorn; Elm; Holly; and Douglas Fir. Spacing: low (3.2m). Canopy openness: 16%. Mean height is 18m. Mean DBH is 30cm.	Species: Holly; Alder Buckthorn; and Ash. Density: moderate Understorey: (sparse) grass (15%), and moss (15%).	Standing dead found ($1.8 \text{ m}^3 \text{ ha}^{-1}$) in a moderate to high decay state. Downed deadwood volume was low ($5.4 \text{ m}^3 \text{ ha}^{-1}$) in a moderate decay state.
#12 E 435480; N 102014 (Hawkhill) *Plantation* (Soil pH 4.7)	Species: Douglas Fir; and Scots Pine. Spacing: moderate (5.3m). Canopy openness of 19%. Mean height is 25m. Mean DBH is 43cm.	Species: Scots Pine and Bramble Density: low. Understorey: fern (65%), grass (20%), and moss (15%).	No standing or fallen dead trees. Downed deadwood volume is low ($1 \text{ m}^3 \text{ ha}^{-1}$) in a moderate to high decay state.
#13 E 435616; N 102208 (Hawkhill) *Semi-ancient* (Soil pH 4.6)	Species: Oak; Beech; and Silver Birch. Spacing: low (4.2m). Canopy openness: 16%. Mean height is 22m. Mean DBH is 46cm.	Species: Holly and Silver Birch. Density: moderate. Understorey: fern (50%) and grass (40%).	High volume of standing dead ($450 \text{ m}^3 \text{ ha}^{-1}$) in a moderate to high decay state. Downed deadwood volume was high ($1202 \text{ m}^3 \text{ ha}^{-1}$) in a moderate decay state.
#14 E 435493; N 102483 (Hawkhill) *Plantation* (Soil pH 5.1)	Species: Douglas Fir; Norway Spruce; Silver Birch. Spacing: low (3m). Canopy openness: 21%. Mean height is 23m. Mean DBH is 34cm.	Species: Beech; Holly; Norway Spruce; Scots Pine; Oak; Silver Birch; and Douglas Fir Density: moderate Understorey: fern (70%) and grass (15%).	No standing or fallen dead trees. Downed deadwood volume is low ($5 \text{ m}^3 \text{ ha}^{-1}$) in a moderate to high decay state.
#15 E 434237; N 103181 (Frame-Heath) *Plantation* (Soil pH 5.5)	Species: Beech; Oak; Douglas Fir and Silver Birch. Spacing: low (3m). Canopy openness: 14.7%. Mean height is 17.5m. Mean DBH is 31cm.	Species: Holly Density: low. Understorey: (sparse) Grass (5), and Fern (5%). *Animal browsing	Standing dead found ($0.3 \text{ m}^3 \text{ ha}^{-1}$) in a moderate decay state. Downed deadwood volume was low ($4 \text{ m}^3 \text{ ha}^{-1}$) in a moderate decay state.
#16 E 434311; N 103473 (Frame-Heath) *Plantation* (Soil pH 5.7)	Species: Scots Pine; Norway Spruce; Oak and Ash. Spacing: is low (4m). Canopy openness: 24%. Mean height is 25m. Man DBH is 40cm.	Species: Holly and Ash, Density: low. Understorey: grass (70%), fern (10%) and moss (10%). *Animal browsing	No standing or fallen dead trees. Downed deadwood volume is moderate to high ($85 \text{ m}^3 \text{ ha}^{-1}$) in a moderate decay state.

Table 5.1 (continued)

Plot location	Tree Species Layer	Regeneration & understorey	Deadwood Amount
#17 E 434148; N 103503 (Frame-Heath) *Plantation* (Soil pH 4.9)	Species: Scots Pine; Beech; and Silver Birch. Spacing: moderate (5m). Canopy openness: 24%. Mean height is 23m. Mean DBH is 44cm.	Species: Silver Birch, Density: low. Understorey: grass (50%) and fern (45%).	No standing dead and few fallen trees. Downed deadwood volume is high (104 m ³ ha ⁻¹) in moderate to high decay state.
#18 E 435143; N 104030 (Frame-Heath) *Plantation* (Soil pH 6.6)	Species: Scots Pine; Norway Spruce; and Oak. Spacing: moderate (5m). Canopy openness: 18%. Mean height is 26m. Mean DBH is 45cm.	Species: Scots Pine and Holly, Density: moderate Understorey: Grass (70%) and Fern (20%). *Animal browsing	Standing dead found (6 m ³ ha ⁻¹) in a moderate decay state. Downed deadwood volume was low (11 m ³ ha ⁻¹) in a moderate decay state.
#19 E 435425; N 103880 (Frame-Heath) *Plantation* (Soil pH 4.5)	Species: Silver Birch; Oak, Norway Spruce; Corsican Pine; and Scots Pine. Spacing: low (3m). Canopy openness: 23%. Mean height is 20.5m. Mean DBH is 29cm.	Species: Scots Pine; and Holly. Density: low. Understorey: fern (60%) and grass (35%). *Animal browsing	Standing dead found (4.5 m ³ ha ⁻¹) in a moderate decay state. Downed deadwood volume was low (8.4 m ³ ha ⁻¹) in a moderate decay state.
#20 E 435310; N 103502 (Frame Wood) *Semi-ancient* (Soil pH 6.7)	Species: Oak; Beech; Hawthorn; Holly; and Silver Birch. Spacing: low (4.5m). Canopy openness: 16%. Mean height is 21.5m; Mean DBH is 48cm.	Species: Beech; and Holly, Density: moderate Understorey: fern (50%) and grass (10%). *Animal browsing	Standing dead found (49 m ³ ha ⁻¹) in a moderate decay state. Downed deadwood volume was high (101 m ³ ha ⁻¹) in a moderate decay state.
#21 E. 435648; N 103012 (Hawkhill) *Plantation* (Soil pH 6.0)	Species: Norway Spruce; Douglas Fir; and Silver Birch. Spacing: low (3m). Canopy openness: 10%. Mean height is 24m. Mean DBH is 27cm.	Species: Holly. Density: moderate. Understorey: grass (30%), moss (40%) and fern (10%). *Animal browsing	Standing dead found (1.9 m ³ ha ⁻¹) in a low to moderate decay state. Downed deadwood volume was moderate (34 m ³ ha ⁻¹) in a moderate decay state.

Table 5.2 – Summarised Descriptions for field plots visited in 2012

Plot location	Tree Species Layer	Regeneration & understorey	Deadwood Amount
#22 E 435165 N 102932 (Frame) *semi-ancient*	Species: Oak, Beech, silver birch, Common Hawthorn and Midland hawthorn. Spacing: 5m Canopy openness: 25-50% Mean height: 12.8m Mean DBH: 35cm	Species: Holly Density: very low Understorey: fern (40%), moss (15%) and grass (25%).	Standing dead found (91.8 m ³ ha ⁻¹) in high decay state. Downed deadwood found (5.6 m ³ ha ⁻¹) in a moderate to high decay state.
#23 E434535; N 103238 (Frame-heath) *plantation*	Species: Hybrid larch and silver birch. Spacing: 2m Canopy openness: 5-10% Mean height: 17m Mean DBH: 18cm	Species: None Density: n/a Understorey: fern (20%), moss (50%) and grass (30%).	Standing dead found (1.1 m ³ ha ⁻¹) in a moderate decay state. Downed deadwood (10.6 m ³ ha ⁻¹) in a low to moderate decay state.
#24 E432686; N 105989 (Denny Wood) *Deciduous woodland*	Species: Oak, Beech, Corsican Pine. Spacing: 2m Canopy openness: 15% Mean height: 19m Mean DBH: 20m	Species: Beech and Holly. Density: very low. Understorey: moss (20%), grass (10%), fern (5%) and unidentified flower (2%).	Standing dead found (1 m ³ ha ⁻¹) in a high decay state. Downed deadwood found (11 m ³ ha ⁻¹) in a high decay state.
#25 E 432700; N 106332 (Denny Wood) *deciduous woodland*	Species: Oak, Beech and holly. Spacing: 2.5m Canopy openness: 15% Mean height: 22m Mean DBH: 38cm	Species: Beech and Holly Density: High Understorey: fern (1%), moss (15%) and grass (25%).	Standing dead found (1.3 m ³ ha ⁻¹) in a moderate to high decay state. Downed deadwood found (21.7 m ³ ha ⁻¹) in a moderate to high decay state
#26 E433456; N 105675 (Denny Wood) *deciduous woodland*	Species: Oak, Beech and Holly. Spacing: 2m Canopy openness: 70% Mean height: 21m Mean DBH: 41cm	Species: Oak, Beech and Holly. Density: Moderate. Understorey: Bracken (50%), Ivy (0.01%), moss (10%) and grass (30%).	Standing dead found (15 m ³ ha ⁻¹) in a moderate to high decay state. Downed deadwood found (9 m ³ ha ⁻¹) in a moderate to high decay state.
#27 E 434562 N 104627 (Denny-Lodge) *Deciduous woodland*	Species: Oak, Beech and Holly. Spacing: 1.8m Canopy openness: 15% Mean height: 25m Mean DBH: 23cm	Species: Holly, Beech, Privet, Whitebeam, Dogwood. Density: Moderate Understorey: Ivy (0.5%) and Moss (5%).	Standing dead found (37 m ³ ha ⁻¹) in a moderate decay state. Downed deadwood found (103 m ³ ha ⁻¹) in a high decay state.
#28 E434716; N 104281 (Denny-Lodge) *Coniferous plantation*	Species: Scots Pine, Douglas Fir and Oak. Spacing: 3m Canopy openness: 25-50% Mean height: 23m Mean DBH: 35cm	Species: Hawthorn, Silver Birch, Douglas fir, Oak and Scots pine. Density: Low Understorey: Foxglove (0.01%), Bracken (0.01%), Nettle (0.01%), grass (6%) and moss (70%).	Standing dead found (0.08 m ³ ha ⁻¹) in a low decay state. Downed deadwood found (23 m ³ ha ⁻¹) in a moderate decay state.
#29 E433379; N 104757 (Denny-Lodge) *Deciduous woodland*	Species: Beech and Oak. Spacing: 2.3m Canopy openness: 20% Mean height: 18m Mean DBH: 30cm	Species: Beech, Hawthorn, holly and Oak. Density: Low Understorey: Bracken (50%), grass (50%) and Moss (5%).	Standing dead found (0.4 m ³ ha ⁻¹) in a low decay state. Downed deadwood found (20 m ³ ha ⁻¹) in a moderate decay state.

Table 5.2 (continued)

Plot location	Tree Species Layer	Regeneration & understorey	Deadwood Amount
#30 E433959; N 101966 (Stockley) *Coniferous/ Deciduous mix*	Species: Douglas fir, Scots pine, Silver Birch, Sweet Chestnut. Spacing: 1.8m Canopy openness: 5% Mean height: 19m Mean DBH: 22m	Species: Silver Birch, Douglas Fir, Holly. Density: low Understorey: Nettle (0.01%), Ivy (0.01%), bracken (1%), grass (0.05) and moss (7%).	Standing dead found (22 m ³ ha ⁻¹) in a moderate decay state. Downed deadwood found (17 m ³ ha ⁻¹) in a moderate deadwood state.
#31 E432859 N 104606 (Stubby-Copse) *Coniferous plantation*	Species: Douglas fir Spacing: 4.4m Canopy openness: 10-20% Mean height: 29% Mean DBH: 40cm	Species: Holly, Beech, Hawthorn and Norway spruce. Density: low Understorey: Nettle (2%), Clover (0.01%), Moss (40%), grass (25%) and bracken (30%).	Standing dead not found. Downed deadwood found (15.6 m ³ ha ⁻¹) in a moderate decay state.
#32 E 433048; N 104205 (Stubby-Copse) * Coniferous/ Deciduous mix*	Species: Scots Pine, Douglas Fir, Oak and Beech. Spacing: 4.5m Canopy openness: 15% Mean height: 25.6m Mean DBH: 48cm	Species: Oak, Douglas Fir, Beech and Holly. Density: Low. Understorey: Bracken (5%), Grass (10%) and moss (5%).	Standing dead found (19 m ³ ha ⁻¹) in a moderate to high decay state. Downed deadwood found (7 m ³ ha ⁻¹) in a moderate to high decay state.
#33 E433123 N 102447 (Irons-Hill) *Coniferous plantation*	Species: Corsican Pine, western hemlock, birch, lawsons cypress. Spacing: 4m Canopy openness: 30% Mean height: 28m Mean DBH: 39cm	Species: western hemlock Density: high Understorey: grass (5%) and moss (30%).	Standing dead found (0.2 m ³ ha ⁻¹) in a high decay state. Downed deadwood found (7 m ³ ha ⁻¹) in a moderate decay state.
#34 E 433692 N 104195 (Denny-Lodge) *Coniferous/ Deciduous mix*	Species: Corsican pine, douglas fir, Norway spruce, Oak and silver birch. Spacing: 3.5m Canopy openness: 30-40% Mean height: 23m Mean DBH: 31cm	Species: Douglas fir, Norway spruce, silver birch, hawthorn and holly. Density: low Understorey: Apple (0.01%), Dog Violet (0.01%), Ivy (0.01%), bracken (50%), Grass (40%), Thistle (1%), Wood spurge (2%) and Rose (0.1%).	Standing dead found (1.2 m ³ ha ⁻¹) in a high decay state. Downed deadwood found (426 m ³ ha ⁻¹) in a moderate to high decay state.
#35 E 434052 N 104683 (Denny-Lodge) *Coniferous Plantation*	Species: Coast redwood, douglas fir, silver birch, beech. Spacing: 3m Canopy openness: 20% Mean height: 24m Mean DBH: 38cm	Species: Coastal redwood, silver birch, Douglas fir, holly and Hawthorn. Density: High Understorey: Ivy (0.01%), grass (20%), woodspurge (0.01%) and moss (20%).	Standing dead found (1.2 m ³ ha ⁻¹) in a moderate to high decay state. Downed deadwood found (29 m ³ ha ⁻¹) in a moderate to high decay state.
#36 E 433308 N 105150 (Denny-Lodge) *Coniferous woodland*	Species: Douglas Fir, Scots pine, silver birch and Oak. Spacing: 4m Canopy openness: 60% Mean height: 25m Mean DBH: 37cm	Species: Holly, Oak, Scots pine, hawthorn. Density: moderate Understorey: Heather (1%), Nettle (1%), Bracken (60%), Grass (30%) and moss (50%).	Standing dead found (0.1 m ³ ha ⁻¹) with a high decay state. Downed deadwood found (15 m ³ ha ⁻¹) with a moderate to high decay state.

Table 5.2 (continued)

Plot location	Tree Species Layer	Regeneration & understorey	Deadwood Amount
#37 E 434339 N 104046 (Denny-Lodge) *Deciduous woodland*	Species: Hawthorn, Beech and oak. Spacing: 3.5m Canopy openness: 20% Mean height: 18m Mean DBH: 38cm	Species: Hawthorn, Beech, Holly and Scots pine. Density: moderate Understorey: Bracken (15%), grass(2%), Ivy(1%), moss (5%) and nettles (1%).	Standing dead not found. Downed deadwood found (24 m ³ ha ⁻¹) with a moderate to high decay state.
#38 E 432837 N 105538 (Denny-Lodge) *Coniferous dominated woodland*	Species: Western Hemlock, Oak and Beech. Spacing: 3m Canopy openness: 20% Mean height: 21m Mean DBH: 30cm	Species: Holly Density: low Understorey: Moss (3%), Bracken (2%) and grass (2%).	Standing dead not found. Downed deadwood found (19 m ³ ha ⁻¹) with a moderate to high decay state.
#39 E 432695 N 106680 (Denny Wood) *Coniferous woodland*	Species: Douglas Fir Spacing: 5m Canopy openness: 40% Mean height: 33m Mean DBH: 39cm	Species: Holly, Oak, Douglas fir, Hawthorn and Lawsons Cypress. Density: moderate Understorey: Blackberry (0.5%), sage (1%), foxglove (1%), clover (2%), grass (60%) and moss (40%).	Standing dead found (0.8 m ³ ha ⁻¹) with a moderate to high decay state. Downed deadwood found (61 m ³ ha ⁻¹) with a moderate to high decay state.
#40 E 433070 N 103862 (Perrywood Haseley) *deciduous woodland*	Species: Oak beech, douglas fir Spacing: 4m Canopy openness: 20% Mean height: 24m Mean DBH: 34cm	Species: Douglas Fir, Beech, Hawthorn, Silver Birch and Oak. Density: moderate Understorey: Grass (80%), moss (15%), bracken (0.01%), unknown (0.01%).	Standing dead not found. Downed deadwood found (43 m ³ ha ⁻¹) with a moderate to high decay state.
#41 E 432632 N 103088 (New Copse) *Coniferous/Deciduous mix*	Species: Beech, Oak, Corsican Pine, Scots pine, Norway Spruce. Spacing: 3.3m Canopy openness: 60% Mean height: 19m Mean DBH: 33cm	Species: Holly, Bramble, Douglas Fir, Beech, Scots pine and Oak. Density: moderate Understorey: Blackberry (0.01%), Ivy (1%), Vetch (0.01%), Grass (20%), Bracken (30%) and moss (25%).	Standing dead not found. Downed deadwood found (31 m ³ ha ⁻¹) with a moderate to high decay state.

A summary of the 39 field plot metrics for the 2010 fieldwork plots is given in Table 5.3, whilst the metrics for the 2012 fieldwork are summarised in Table 5.4. It should be noted that canopy openness measurements were only available for field data collected in 2010. There are a number of trends evident within this dataset. In terms of the forest structure, generally, deciduous field plots exhibit a smaller stem number, contain a higher proportion of native tree species, have higher DBH values, higher DBH standard deviation values, a higher percentage of large trees, lower height to the first live branch, larger average and total crown horizontal areas than their coniferous or mixed counterparts for both 2010 and 2012. There

are a few examples of deciduous plots containing a much higher stem density, for example plot 24, which contains 53 stems, but also has much lower mean DBH (19.52cm) and average crown horizontal area (29.48m²) values, whereas plot 39, located in a coniferous stand, contains 16 stems, has a relatively high mean DBH (39.61cm) and a higher average crown horizontal area (50.96m²). On average, coniferous and deciduous plots contain similar numbers of trees or tree species. Additionally, there is no clear difference between the numbers of canopy layers between the two broad tree types. Estimates of canopy openness and average tree spacing also vary between all plot sites and broad tree types.

In terms of deadwood, the majority of deciduous plot sites contain a higher volume of standing deadwood, although some plots contain none at all, such as with plots 37 and 40. Many of the plots surveyed in 2010 within coniferous stands did not contain much in the way of standing deadwood. For coniferous plots the decay class for standing deadwood varied from low (0.3) to high (1.0), the majority occupying the range of 0.55-0.70. For deciduous plots the decay class showed less variability between plots and ranged from 0.59-0.98, with the majority occupying the range of 0.67-0.83. Again, the larger values for downed deadwood (DDW) volume are mostly encountered within deciduous plots; however there is a large range of 0.33-14.62m³. Plots within coniferous stands are typically more similar, varying by $\pm 2\text{m}^3$. There are a number of plots where the DDW volume for coniferous plots is very high, the most dramatic example being plot 34, which is a mixed (about one third deciduous) but predominantly coniferous plot, with approximately one-third of the population being deciduous trees. The DDW decay class for coniferous plots ranges from 0.48-0.9 and is quite variable between plots. For deciduous plots, the range is less variable from 0.69-0.95.

In terms of understorey composition, many of the coniferous and deciduous plots have no saplings whatsoever (17 plots). Generally, sapling population is very low overall, and is typically below 10 per plot. There are two exceptions however, plot 14 (coniferous) and plot 27 (deciduous) which contain drastically higher populations. The majority of saplings were of native species. The maximum number of sapling species was 4, but the average was 2. There was a much higher proportion of seedlings discovered through all but one of the field plots visited, where the values ranged from 9 to 936. There was again a great deal of variability in population numbers, but deciduous plot sites typically recorded the highest values. For coniferous plots there was a greater disparity in seedling counts, although the highest population numbers were recorded for this woodland type. The majority of seedlings were of

native species. Species number ranged from 1 to 17, the average value was 4, and was variable between both coniferous and deciduous woodland types. A Shannon-Wiener (SH) diversity index was calculated for native seedlings for each plot, and was extremely variable between plots, ranging from 0.2-1.41.

The number of vascular plant species within the plots varied between 2 and 8. Coniferous plots generally recorded higher values, but included the most variability. The number of bryophyte species encountered within each plot varied between 5 and 25, and was highly variable between plots, but seemed to achieve higher values with higher values of canopy openness. The estimate of percentage bare soil was again highly variable between plots.

To examine the similarity between the field-plot data collected in 2010 and 2012 for each of the 37 field attribute populations, summary statistics (mean, standard deviation and range) were calculated. These populations were stratified by (i) the total number of field plots (see Table 5.5), (ii) those of predominantly coniferous species (see Table 5.6), (iii) those of predominantly deciduous species (see Table 5.7), and finally (iv) those of mixed species (see Table 5.8).

The first of these, concerning data collected across all field-plots (Table 5.5), indicates on average, similar values for the majority of field metrics. There are exceptions however, for example a slightly higher mean stem density and total number of seedlings for data collected in 2012 over that collected in 2010. This pattern is again repeated for the coniferous dominated subset (Table 5.6). For both the predominantly deciduous subset (Table 5.7) and those of mixed species (Table 5.8), the stem density is more similar between 2010 and 2012 data; however the number of seedlings is much higher for 2012.

Table 5.3 – 2010 Field metrics summary

Site No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Mean	STD	Range
Class	Mixed	Mixed	Conf.	Decid.	Conf.	Decid.	Conf.	Decid.	Conf.	Decid.	Conf.	Decid.	Conf.	Decid.	Conf.	Decid.	Conf.	Conf.	Mixed	Decid.	Conf.			
Canopy openness (%)	20.59	17.24	19.59	21.01	22.81	12.97	26.84	22.76	15.68	14.67	15.8	18.89	15.72	21.33	14.75	23.87	-	17.72	22.8	16.21	10.55	2.10	0.70	2.00
No. Canopy layers	2	2	3	3	1	1	2	2	2	1	2	1	2	2	3	2	2	2	2	3	3	2.10	0.70	2.00
Number of trees per 30x30m	27	31	19	11	30	7	17	36	11	20	36	20	17	25	44	24	26	21	32	16	52	24.86	11.12	45.00
Number of native trees per 30x30m	27	2	15	11	14	7	7	6	11	20	28	2	17	4	34	22	26	15	30	16	15	15.67	9.46	32.00
Percentage of native trees	100.00	6.45	78.95	100.00	46.67	100.00	41.18	16.67	100.00	100.00	77.78	100.00	16.00	77.27	91.67	100.00	71.43	93.75	100.00	28.55	69.36	35.28	95.55	95.55
No. of tree species per 30x30m	2	2	4	3	3	1	4	4	3	3	7	2	3	4	4	6	3	3	5	5	3	3.52	1.40	6.00
Average tree spacing (m)	2.67	3.86	4.65	4.73	3.56	8.45	4.34	2.90	5.29	2.65	3.16	5.19	4.12	3.39	3.02	3.90	4.78	5.09	2.78	4.32	2.97	4.09	1.33	5.80
STDxv. Of tree spacing (m)	1.45	1.36	1.84	3.08	1.24	2.67	1.51	1.21	1.71	1.86	2.12	1.56	2.60	2.00	1.19	1.82	1.15	1.04	1.87	2.65	0.56	1.76	0.59	2.12
Mean DBH (cm)	30.66	34.28	35.85	57.21	34.37	61.84	40.76	33.19	48.21	49.85	30.42	42.78	45.61	33.93	30.85	39.46	43.74	44.62	29.18	47.59	27.45	40.06	9.41	34.39
STDxv. of DBH (cm)	12.86	7.30	11.20	35.50	7.62	18.39	12.57	13.73	21.68	28.88	15.72	8.50	23.44	15.01	15.40	13.24	9.40	12.97	14.13	38.58	13.00	16.62	8.57	31.21
Basal area within 30x30m	2.59	2.99	2.10	3.82	2.92	2.26	2.36	3.63	2.38	5.24	3.02	2.98	3.67	2.69	4.09	3.25	4.08	3.55	2.71	4.60	3.98	3.28	0.83	3.14
Percentage big trees (DBH40x>80cm)	16.67	22.58	36.84	27.27	23.33	71.43	58.82	38.89	72.73	40.00	18.18	60.00	38.89	56.00	26.19	66.67	69.23	61.91	21.21	31.25	26.52	42.14	19.48	56.06
Percentage very big trees (DBH>80cm)	0.00	0.00	0.00	27.27	0.00	14.29	0.00	0.00	0.00	15.00	0.00	0.00	11.11	0.00	0.00	0.00	0.00	0.00	0.00	18.75	0.00	4.12	8.04	27.27
Percentage big trees (DBH>40cm)	16.67	22.58	36.84	27.27	23.33	71.43	58.82	38.89	72.73	40.00	18.18	60.00	50.00	56.00	26.19	66.67	69.23	61.91	21.21	50.00	26.52	46.26	20.44	66.05
Loresys Mean Height (m)	18.22	18.35	17.64	17.49	20.53	24.25	22.02	16.90	22.54	24.63	17.86	23.60	18.99	23.21	17.51	25.07	23.48	27.01	20.51	21.57	23.61	21.19	3.05	10.11
Shannon Index (by stem count)	0.16	0.24	1.19	0.99	0.88	0.00	1.15	1.13	0.93	0.85	1.14	0.33	0.83	0.94	1.07	1.16	0.99	0.84	1.13	1.47	1.07	0.86	0.39	1.47
Shannon Index (by stem count)	0.07	0.12	1.00	0.67	0.57	0.00	0.67	0.65	0.64	0.56	0.75	0.00	0.54	0.52	0.55	0.28	0.34	0.51	0.59	0.80	0.66	0.50	0.27	1.00
Shannon Index (by BA)	0.42	0.35	1.02	0.24	0.76	0.00	0.85	1.02	0.75	0.57	0.47	0.49	0.48	0.30	1.14	0.67	0.35	0.43	0.33	0.54	0.70	0.57	0.29	1.14
Simpson Index (by BA)	0.42	0.20	0.59	0.95	1.00	0.00	0.74	0.76	0.88	0.40	0.89	0.47	0.37	0.19	0.87	0.39	0.27	0.72	1.18	0.42	0.53	0.58	0.31	1.11
Average height of the first live branch (m)	8.41	8.78	6.77	3.47	10.78	7.23	7.02	7.02	3.30	6.84	4.57	13.94	5.55	8.94	6.61	10.47	13.41	12.47	7.03	3.58	10.74	8.00	3.20	10.66
STDxv. of the height of the first live branch (m)	3.26	2.30	2.94	1.69	3.24	1.42	4.78	2.95	1.14	4.85	2.76	1.54	2.76	5.46	4.01	5.62	4.06	4.83	3.92	2.10	3.86	3.31	1.34	4.44
Average crown horizontal area (m2)	25.96	48.91	56.75	107.68	45.30	119.91	51.44	26.75	104.32	103.88	50.03	51.09	52.01	23.61	46.78	48.65	56.95	48.53	30.54	82.27	26.84	57.52	28.96	96.30
STDxv. Crown horizontal area (m2)	21.87	27.28	35.22	109.10	1.39	74.02	1.30	14.08	46.65	55.97	45.94	27.89	33.64	11.44	30.73	20.77	23.54	29.22	26.94	74.76	23.96	35.04	25.89	107.80
Total Crown horizontal area (m2)	778.73	1516.29	1078.23	1184.46	1358.97	839.36	874.46	963.10	1147.55	2073.68	1650.85	1021.82	958.19	590.17	2198.75	1167.55	1480.76	1019.16	1007.96	1316.37	1395.61	1219.05	402.72	1608.58
Standing dead wood volume (m3)	0.00	0.00	0.00	4.24	0.07	4.26	0.00	0.28	10.12	6.92	0.16	0.80	8.90	0.00	0.03	0.00	0.00	0.57	0.41	4.45	0.17	1.93	3.22	10.11
Standing dead wood decay class	-	-	-	0.78	0.73	0.78	-	0.67	0.92	0.83	0.96	-	0.67	-	-	-	-	0.67	0.67	0.67	0.42	0.73	0.14	0.54
Downed dead wood volume (m3)	0.91	1.25	0.16	0.82	0.78	0.93	1.47	0.30	8.78	14.62	0.49	0.09	12.02	0.48	0.33	7.61	9.43	1.00	0.76	9.13	3.05	3.54	4.58	14.53
Downed dead wood decay class	0.58	0.71	0.52	0.91	0.84	0.95	0.90	0.52	0.87	0.75	0.88	0.83	0.78	0.69	0.93	0.73	0.89	0.69	0.83	0.75	0.85	0.78	0.13	0.43
No. saplings per 30x30m	5.00	0.00	0.00	0.00	0.00	7.00	0.00	1.00	1.00	0.00	14.00	4.00	5.00	42.00	3.00	0.00	0.00	0.00	3.00	4.00	0.00	4.33	9.31	42.00
No. native saplings per 30x30m	5.00	0.00	0.00	0.00	0.00	7.00	0.00	1.00	1.00	0.00	14.00	4.00	5.00	34.00	3.00	0.00	0.00	0.00	3.00	4.00	0.00	3.86	7.72	34.00
Percentage of native saplings per 30x30m	100.00	0.00	0.00	0.00	0.00	100.00	0.00	100.00	100.00	0.00	100.00	100.00	80.95	100.00	100.00	0.00	0.00	0.00	60.00	100.00	0.00	49.57	49.27	100.00
No. of sapling species per 30x30m	2.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	1.00	0.00	3.00	1.00	2.00	4.00	1.00	0.00	0.00	0.00	2.00	3.00	0.00	1.00	1.22	4.00
No. seedlings per 30x30m	18.00	18.00	18.00	18.00	27.00	45.00	864.00	54.00	333.00	342.00	405.00	18.00	162.00	216.00	18.00	27.00	9.00	198.00	18.00	180.00	63.00	145.29	206.44	855.00
No. native seedlings per 30x30m	18.00	18.00	18.00	18.00	27.00	45.00	144.00	45.00	333.00	342.00	405.00	18.00	162.00	189.00	18.00	27.00	9.00	198.00	18.00	180.00	63.00	109.29	113.94	396.00
Percentage of native seedlings per 30x30m	100.00	100.00	100.00	100.00	100.00	100.00	16.67	83.33	100.00	100.00	100.00	100.00	100.00	87.50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	94.64	88.41	83.33
No. of seedling species per 30x30m	1.00	1.00	1.00	2.00	1.00	1.00	4.00	2.00	3.00	2.00	3.00	2.00	2.00	7.00	1.00	2.00	1.00	2.00	2.00	2.00	1.00	2.05	1.40	6.00
Seedlings Shannon Index for native species	0.00	0.00	0.00	0.00	0.00	0.00	0.56	0.00	0.66	0.39	0.64	0.69	0.45	1.27	0.00	0.64	0.00	0.47	0.00	0.20	0.00	0.30	0.38	1.27
No. Vascular Species	2.00	4.00	3.00	3.00	2.00	6.00	4.00	4.00	2.00	5.00	2.00	4.00	6.00	6.00	3.00	3.00	4.00	4.00	4.00	6.00	4.00	3.86	1.35	4.00
No. Biotrophic Species	5.00	15.00	10.00	20.00	10.00	15.00	10.00	10.00	25.00	25.00	15.00	15.00	20.00	15.00	25.00	15.00	10.00	10.00	15.00	5.00	15.00	14.52	5.90	20.00
Percentage of bare soil within 30x30m	0.50	0.00	10.00	30.00	5.00	50.00	5.00	30.00	0.00	80.00	5.00	0.00	5.00	30.00	90.00	10.00	5.00	5.00	10.00	50.00	35.00	21.69	26.45	90.00

Table 5.4– 2012 Field metrics summary

Site No.	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	Mean	STD	Range
Class	Decid.	Conif.	Decid.	Decid.	Decid.	Decid.	Conif.	Decid.	Conif.	Mixed	Conif.	Mixed	Conif.	Conif.	Decid.	Mixed	Conif.	Decid.	Mixed	Decid.			
Canopy openness (%)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
No. Canopy layers	2	2	2	2	1	3	1	2	2	1	1	1	3	3	2	2	3	2	3	2	2.00	0.73	2.00
Number of trees per 30x30m	20	71	53	21	17	33	35	30	90	24	20	28	28	33	25	23	36	16	17	24	32.20	18.86	74.00
Number of native trees per 30x30m	20	25	52	21	16	33	12	30	42	0	13	3	13	4	24	23	13	16	15	22	19.85	12.66	52.00
Percentage of native trees	100.00	35.21	98.11	100.00	94.12	100.00	34.29	100.00	46.67	0.00	65.00	10.71	46.43	12.12	96.00	100.00	36.11	100.00	86.24	91.67	67.73	36.03	100.00
No. of tree species per 30x30m	5	2	3	3	5	3	3	2	5	1	4	4	5	4	4	4	3	1	3	5	3.45	1.28	4.00
Average tree spacing (m)	-	2.38	1.97	2.29	1.88	1.87	3.10	2.27	1.84	4.43	4.50	3.99	3.51	3.04	4.25	3.56	3.00	5.17	4.41	3.32	3.20	1.05	3.32
STD of tree spacing (m)	-	1.21	1.37	1.59	1.57	0.94	0.87	1.34	0.83	1.09	1.29	0.94	1.66	1.47	1.13	1.37	1.05	1.65	1.46	1.33	1.27	0.27	0.83
Mean DBH (cm)	35.56	18.72	19.62	38.07	40.63	22.76	35.43	30.23	22.88	41.36	47.86	38.54	31.44	37.96	37.23	37.92	30.46	39.61	34.42	33.46	33.71	7.72	29.14
STD of DBH (cm)	20.82	7.37	12.60	22.13	20.69	23.86	7.87	17.06	13.60	6.91	15.58	10.86	15.45	22.48	11.53	14.95	12.47	15.87	19.59	55.67	17.36	10.36	48.76
Basal area within 30x30m	2.90	2.28	2.35	3.46	2.83	3.63	3.62	2.91	5.27	3.31	3.96	3.52	2.68	5.00	2.97	2.98	4.23	2.27	2.06	2.79	3.26	0.86	3.21
Percentage big trees (DBH 40x-80cm)	35.00	0.00	0.68	38.10	52.94	27.27	25.71	23.33	8.89	58.33	70.00	46.43	32.14	21.21	52.00	43.48	19.44	56.25	17.65	33.33	33.11	19.31	70.00
Percentage very big trees (DBH >80cm)	0.00	0.00	0.00	0.00	5.88	0.00	0.00	0.00	1.11	0.00	0.00	0.00	0.00	6.06	0.00	0.00	0.00	0.00	5.88	0.00	0.95	2.17	6.06
Loeys Mean Height (m)	35.00	0.00	7.55	38.10	58.62	27.27	25.71	23.33	10.00	58.33	70.00	46.43	32.14	27.27	52.00	43.48	19.44	56.25	23.53	33.33	34.40	18.66	70.00
Shannon Index (by stem count)	1.39	0.65	0.46	0.98	1.15	0.96	0.74	0.54	0.98	0.00	1.35	0.74	1.43	1.02	0.82	0.78	0.90	0.00	0.87	1.41	0.86	0.41	1.43
Shannon Index (by stem count)	0.90	0.39	0.86	0.81	0.97	0.77	0.75	0.68	1.29	0.00	1.24	0.30	1.15	0.67	0.58	0.50	1.09	0.00	0.70	0.99	0.73	0.36	1.29
Shannon Index (by BA)	0.75	0.46	0.24	0.61	0.63	0.59	0.48	0.37	0.57	0.00	0.77	0.38	0.76	0.60	0.46	0.50	0.54	0.00	0.54	0.50	0.49	0.21	0.77
Shannon Index (by BA)	0.77	0.41	0.86	0.75	0.80	0.71	0.66	0.73	0.84	0.00	0.93	0.20	1.01	0.51	0.41	1.60	0.97	0.00	2.89	0.97	0.80	0.62	2.89
Average height of the first live branch (m)	2.83	8.89	7.18	6.46	6.53	5.54	11.07	5.37	10.33	12.29	9.01	13.46	7.13	7.18	11.51	5.09	2.63	11.53	5.26	6.30	7.78	3.10	10.84
STD of the height of the first live branch (m)	1.62	3.05	-	4.03	4.53	5.92	2.39	3.41	3.40	2.56	4.22	4.07	6.11	3.25	5.19	3.21	0.95	5.21	3.27	5.89	3.80	1.44	5.16
Average crown horizontal area (m ²)	48.67	13.03	29.48	100.11	59.68	34.39	32.54	49.65	-	41.66	87.13	23.52	41.07	33.48	32.32	64.86	43.59	50.96	91.14	66.54	49.67	23.55	87.08
STD of crown horizontal area (m ²)	973.32	925.16	-	2102.26	1014.56	1134.96	1138.78	1489.55	-	999.74	1742.62	658.59	1149.84	1104.84	808.06	1491.88	1569.10	815.33	1546.31	1597.06	1236.54	382.30	1443.67
Standing dead wood volume (m ³)	8.27	0.10	0.09	0.12	1.36	3.42	0.01	0.04	1.99	0.00	1.74	0.02	0.11	0.11	0.01	0.00	0.00	0.07	0.00	0.00	0.74	0.18	0.67
Standing dead wood decay class	0.78	0.67	0.78	0.78	0.67	0.59	0.33	1.00	0.55	-	0.67	1.00	0.83	0.73	1.00	-	-	0.67	-	-	4.44	8.29	38.21
Downed dead wood volume (m ³)	6.21	2.44	1.62	0.22	0.85	9.32	2.13	1.81	2.70	1.41	0.62	0.67	38.42	2.67	1.43	2.21	1.72	5.50	3.95	2.82	0.88	0.11	0.47
Downed dead wood decay class	0.72	0.50	0.95	0.70	0.77	0.85	0.71	0.69	0.51	0.48	0.71	0.52	0.73	0.65	0.71	0.73	0.69	0.66	0.71	0.69	9.85	24.74	108.00
No. saplings per 30x30m	0.00	0.00	0.00	0.00	0.00	38.00	8.00	2.00	15.00	1.00	2.00	0.00	108.00	4.00	1.00	7.00	0.00	0.00	10.00	1.00	2.90	8.45	38.00
No. native saplings per 30x30m	0.00	0.00	0.00	0.00	0.00	38.00	2.00	2.00	0.00	0.00	0.00	0.00	0.00	2.00	1.00	7.00	0.00	0.00	1.00	1.00	0.00	0.97	3.00
Percentage of native saplings per 30x30m	0.00	0.00	0.00	0.00	0.00	100.00	25.00	100.00	26.66	0.00	0.00	0.00	0.00	50.00	100.00	100.00	0.00	0.00	10.00	100.00	30.58	43.06	100.00
No. of sapling species per 30x30m	0.00	0.00	0.00	0.00	0.00	2.00	2.00	1.00	2.00	1.00	1.00	0.00	2.00	3.00	1.00	2.00	0.00	0.00	2.00	1.00	1.00	1.00	3.00
No. seedlings per 30x30m	135.00	0.00	45.00	702.00	798.00	996.00	117.00	45.00	81.00	9.00	207.00	404.00	405.00	729.00	522.00	405.00	36.00	117.00	155.00	351.00	339.70	297.62	936.00
No. native seedlings per 30x30m	135.00	0.00	45.00	702.00	702.00	996.00	63.00	45.00	81.00	90.00	99.00	4.00	54.00	54.00	522.00	405.00	36.00	117.00	155.00	351.00	230.60	278.95	936.00
Percentage of native seedlings per 30x30m	100.00	0.00	100.00	100.00	97.56	100.00	53.85	100.00	100.00	83.33	47.83	1.00	13.33	7.41	100.00	100.00	100.00	81.25	22.08	79.59	69.36	39.13	100.00
No. of seedling species per 30x30m	1.00	0.00	2.00	3.00	3.00	6.00	13.00	4.00	2.00	6.00	6.00	3.00	6.00	5.00	4.00	5.00	1.00	5.00	5.00	8.00	4.40	2.89	13.00
Seedling Shannon Index for native species	0.00	0.00	0.67	1.00	0.55	1.12	0.88	1.33	0.35	1.58	1.41	0.69	0.67	1.05	1.15	0.73	0.00	1.24	1.04	1.61	0.84	0.51	1.61
No. Vascular Species	3.00	3.00	4.00	3.00	5.00	3.00	5.00	4.00	6.00	5.00	3.00	2.00	8.00	4.00	6.00	4.00	3.00	8.00	4.00	8.00	4.55	1.82	6.00
No. Bryophyte Species	20.00	20.00	25.00	25.00	15.00	15.00	20.00	15.00	15.00	15.00	15.00	15.00	20.00	20.00	15.00	20.00	20.00	15.00	25.00	20.00	18.50	3.66	10.00
Percentage of bare soil within 30x30m	5.00	0.00	0.00	0.00	45.00	95.00	20.00	90.00	90.00	5.00	60.00	80.00	5.00	70.00	5.00	80.00	90.00	5.00	75.00	50.00	43.50	38.01	95.00

Table 5.5 – Summary values (mean, standard deviation and range) for all field plots for each attribute in 2010 and 2012 data.

	2010 all			2012 all		
	Mean	STD	Range	Mean	STD	Range
No. Canopy Layers	2.10	0.70	2.00	2.00	0.73	2.00
Number of trees per 30x30m	24.86	11.12	45.00	32.20	18.86	74.00
Number of native trees per 30x30m	15.67	9.46	32.00	19.85	12.66	52.00
Percentage of native trees	69.36	35.28	93.55	67.73	36.03	100.00
No. of tree species per 30x30m	3.52	1.40	6.00	3.45	1.28	4.00
Average tree spacing (m)	4.09	1.33	5.80	3.20	1.05	3.32
STDev. Of tree spacing (m)	1.76	0.59	2.12	1.27	0.27	0.83
Mean DBH (cm)	40.06	9.41	34.39	33.71	7.72	29.14
STDev. of DBH (cm)	16.62	8.57	31.28	17.36	10.36	48.76
Basal area within 30x30m	3.28	0.83	3.14	3.26	0.86	3.21
Percentage big trees (DBH 40>x<80cm)	42.14	19.48	56.06	33.11	19.31	70.00
Percentage very big trees (DBH >80cm)	4.12	8.04	27.27	0.95	2.17	6.06
Percentage big trees (DBH >40cm)	46.26	20.44	69.05	34.40	18.66	70.00
Loreys Mean Height (m)	21.19	3.05	10.11	22.40	4.70	20.30
Shannon Index (by stem count)	0.86	0.39	1.47	0.86	0.41	1.43
Simpson Index (by stem count)	0.50	0.27	1.00	0.73	0.36	1.29
Shannon Index (by BA)	0.57	0.29	1.14	0.49	0.21	0.77
Simpson Index (by BA)	0.58	0.31	1.18	0.80	0.62	2.89
Average height of the first live branch (m)	8.00	3.20	10.64	7.78	3.10	10.84
STDev. of the height of the first live branch (m)	3.31	1.34	4.48	3.80	1.44	5.16
Average crown horizontal area (m2)	57.52	28.96	96.30	49.67	23.55	87.08
STDev. Crown horizontal area (m2)	35.04	25.89	107.80	30.82	15.99	57.19
Total Crown horizontal area (m2)	1219.05	402.72	1608.58	1236.94	382.30	1443.67
Standing dead wood volume (m3)	1.93	3.22	10.12	0.87	1.97	8.27
Standing dead wood decay class	0.73	0.14	0.54	0.74	0.18	0.67
Downed dead wood volume (m3)	3.54	4.58	14.53	4.44	8.29	38.21
Downed dead wood decay class	0.78	0.13	0.43	0.68	0.11	0.47
No. saplings per 30x30m	4.33	9.31	42.00	9.85	24.74	108.00
No. native saplings per 30x30m	3.86	7.72	34.00	2.90	8.45	38.00
Percentage of native saplings per 30x30m	49.57	49.27	100.00	30.58	43.06	100.00
No. of sapling species per 30x30m	1.00	1.22	4.00	1.00	0.97	3.00
No. seedlings per 30x30m	145.29	206.44	855.00	339.70	297.62	936.00
No. native seedlings per 30x30m	109.29	123.94	396.00	230.60	278.95	936.00
Percentage of native seedlings per 30x30m	94.64	18.41	83.33	69.36	39.13	100.00
No. of seedling species per 30x30m	2.05	1.40	6.00	4.40	2.89	13.00
Seedlings Shannon Index for native species	0.30	0.38	1.27	0.84	0.51	1.61
No. Vascular Species	3.86	1.35	4.00	4.55	1.82	6.00
No. Bryophyte Species	14.52	5.90	20.00	18.50	3.66	10.00
Percentage of bare soil within 30x30m	21.69	26.45	90.00	43.50	38.01	95.00

Table 5.6 – Summary values (mean, standard deviation and range) for the coniferous field plots for each attribute in 2010 and 2012 data.

Class	2010 Conif.			2012 Conif.		
	Mean	STD	Range	Mean	STD	Range
Site No.						
No. Canopy Layers	2.00	0.67	2.00	1.75	0.71	2.00
Number of trees per 30x30m	27.00	10.42	35.00	40.25	26.01	74.00
Number of native trees per 30x30m	12.60	7.81	24.00	15.75	14.17	42.00
Percentage of native trees	50.14	33.23	90.00	41.87	37.88	100.00
No. of tree species per 30x30m	3.60	1.07	4.00	3.00	1.51	4.00
Average tree spacing (m)	4.08	0.85	2.29	3.53	1.12	3.32
STDev. Of tree spacing (m)	1.44	0.37	1.04	1.15	0.29	0.82
Mean DBH (cm)	37.57	5.52	17.17	33.96	8.38	22.64
STDev. of DBH (cm)	11.72	2.45	7.39	12.06	5.25	15.57
Basal area within 30x30m	3.15	0.66	1.98	3.53	1.12	3.00
Percentage big trees (DBH 40>x<80cm)	49.86	16.81	45.90	33.60	22.62	58.33
Percentage very big trees (DBH >80cm)	0.00	0.00	0.00	0.90	2.12	6.06
Percentage big trees (DBH >40cm)	49.86	16.81	45.90	34.50	22.07	58.33
Loreys Mean Height (m)	22.31	3.15	10.11	24.86	5.26	16.27
Shannon Index (by stem count)	0.93	0.28	0.86	0.62	0.40	1.02
Simpson Index (by stem count)	0.52	0.27	1.00	0.50	0.43	1.29
Shannon Index (by BA)	0.66	0.26	0.72	0.37	0.24	0.60
Simpson Index (by BA)	0.57	0.25	0.81	0.38	0.30	0.84
Average height of the first live branch (m)	10.18	2.63	7.17	10.78	1.98	6.28
STDev. of the height of the first live branch (m)	3.93	1.29	4.08	3.64	1.09	2.82
Average crown horizontal area (m2)	43.59	12.85	33.34	32.50	12.15	37.93
STDev. Crown horizontal area (m2)	18.88	11.54	33.92	16.89	7.65	19.48
Total Crown horizontal area (m2)	1094.98	268.00	890.59	921.50	173.31	480.19
Standing dead wood volume (m3)	0.11	0.19	0.57	0.29	0.69	1.99
Standing dead wood decay class	0.62	0.14	0.31	0.71	0.24	0.67
Downed dead wood volume (m3)	2.44	3.35	9.34	2.37	1.45	4.84
Downed dead wood decay class	0.75	0.14	0.38	0.59	0.10	0.23
No. saplings per 30x30m	4.70	13.17	42.00	3.63	5.37	15.00
No. native saplings per 30x30m	3.90	10.65	34.00	1.13	1.46	4.00
Percentage of native saplings per 30x30m	28.10	45.53	100.00	25.21	35.36	100.00
No. of sapling species per 30x30m	0.60	1.26	4.00	1.13	1.13	3.00
No. seedlings per 30x30m	149.40	262.09	855.00	250.75	268.65	729.00
No. native seedlings per 30x30m	73.80	74.07	189.00	116.38	168.77	522.00
Percentage of native seedlings per 30x30m	88.75	26.06	83.33	53.35	44.28	100.00
No. of seedling species per 30x30m	2.30	1.89	6.00	4.75	3.85	13.00
Seedlings Shannon Index for native species	0.37	0.44	1.27	0.85	0.51	1.58
No. Vascular Species	3.80	1.03	4.00	4.88	1.89	6.00
No. Bryophyte Species	12.00	2.58	5.00	16.88	2.59	5.00
Percentage of bare soil within 30x30m	13.50	12.92	35.00	34.38	38.59	90.00

Table 5.7 – Summary values (mean, standard deviation and range) for the deciduous field plots for each attribute in 2010 and 2012 data.

Class	2010 Decid.			2012 Decid.		
	Mean	STD	Range	Mean	STD	Range
Site No.						
No. Canopy Layers	2.13	0.83	2.00	2.13	0.64	2.00
Number of trees per 30x30m	20.25	13.02	37.00	26.75	12.08	36.00
Number of native trees per 30x30m	18.00	9.13	27.00	26.25	12.14	37.00
Percentage of native trees	94.38	10.40	22.73	97.56	4.29	11.76
No. of tree species per 30x30m	3.63	1.77	6.00	3.50	1.07	3.00
Average tree spacing (m)	4.47	1.85	5.80	2.61	0.99	2.53
STDev. Of tree spacing (m)	2.24	0.62	1.89	1.38	0.22	0.65
Mean DBH (cm)	46.45	11.15	31.42	32.42	7.66	21.31
STDev. of DBH (cm)	24.70	8.82	23.18	18.95	3.82	11.36
Basal area within 30x30m	3.64	1.04	2.98	2.90	0.51	1.56
Percentage big trees (DBH 40>x<80cm)	40.74	20.57	54.55	29.81	16.31	52.26
Percentage very big trees (DBH >80cm)	10.80	10.10	27.27	1.47	2.72	5.88
Percentage big trees (DBH >40cm)	51.55	22.03	67.54	32.13	15.42	51.28
Loreys Mean Height (m)	20.61	3.01	7.14	20.03	3.95	11.96
Shannon Index (by stem count)	0.91	0.42	1.47	0.89	0.30	0.93
Simpson Index (by stem count)	0.56	0.25	0.80	0.77	0.15	0.47
Shannon Index (by BA)	0.52	0.34	1.14	0.53	0.16	0.52
Simpson Index (by BA)	0.60	0.35	0.95	1.14	0.77	2.18
Average height of the first live branch (m)	4.85	1.50	3.54	5.53	1.32	4.36
STDev. of the height of the first live branch (m)	2.59	1.29	3.71	3.71	1.33	4.30
Average crown horizontal area (m2)	83.34	29.79	73.13	59.75	25.14	70.63
STDev. Crown horizontal area (m2)	58.88	26.06	78.37	41.58	15.98	46.45
Total Crown horizontal area (m2)	1418.40	506.91	1359.39	1393.69	394.23	1128.94
Standing dead wood volume (m3)	4.89	3.68	10.09	1.66	2.92	8.27
Standing dead wood decay class	0.79	0.11	0.29	0.77	0.14	0.41
Downed dead wood volume (m3)	5.89	5.89	14.29	3.27	3.09	9.10
Downed dead wood decay class	0.85	0.08	0.20	0.76	0.09	0.26
No. saplings per 30x30m	4.25	4.65	14.00	7.13	13.04	38.00
No. native saplings per 30x30m	4.25	4.65	14.00	6.00	13.15	38.00
Percentage of native saplings per 30x30m	75.00	46.29	100.00	38.75	50.83	100.00
No. of sapling species per 30x30m	1.38	1.19	3.00	0.88	0.99	2.00
No. seedlings per 30x30m	187.88	156.21	387.00	462.38	352.49	891.00
No. native seedlings per 30x30m	187.88	156.21	387.00	392.63	351.28	891.00
Percentage of native seedlings per 30x30m	100.00	0.00	0.00	89.95	27.44	77.92
No. of seedling species per 30x30m	2.00	0.76	2.00	3.63	1.69	5.00
Seedlings Shannon Index for native species	0.32	0.32	0.86	0.78	0.44	1.33
No. Vascular Species	4.13	1.81	4.00	3.75	0.71	2.00
No. Bryophyte Species	18.75	6.94	20.00	20.00	4.63	10.00
Percentage of bare soil within 30x30m	38.75	34.72	90.00	48.75	41.73	95.00

Table 5.8 – Summary values (mean, standard deviation and range) for the mixed field plots for each attribute in 2010 and 2012 data.

	2010 mixed			2012 mixed		
	Mean	STD	Range	Mean	STD	Range
No. Canopy Layers	2.33	0.58	1.00	2.25	0.96	2.00
Number of trees per 30x30m	30.00	2.65	5.00	27.00	6.83	16.00
Number of native trees per 30x30m	19.67	15.37	28.00	15.25	4.50	9.00
Percentage of native trees	66.73	52.30	93.55	59.80	24.38	55.56
No. of tree species per 30x30m	3.00	1.73	3.00	4.25	0.96	2.00
Average tree spacing (m)	3.10	0.66	1.19	3.58	0.65	1.50
STDev. Of tree spacing (m)	1.56	0.27	0.51	1.33	0.25	0.60
Mean DBH (cm)	31.37	2.62	5.10	35.81	8.13	17.39
STDev. of DBH (cm)	11.43	3.63	6.83	24.79	20.63	43.20
Basal area within 30x30m	2.76	0.21	0.40	3.42	0.80	1.56
Percentage big trees (DBH 40>x<80cm)	20.15	3.09	5.91	38.73	21.77	50.56
Percentage very big trees (DBH >80cm)	0.00	0.00	0.00	0.00	0.00	0.00
Percentage big trees (DBH >40cm)	20.15	3.09	5.91	38.73	21.77	50.56
Loreys Mean Height (m)	19.03	1.29	2.29	22.26	2.79	6.65
Shannon Index (by stem count)	0.51	0.54	0.97	1.27	0.25	0.53
Simpson Index (by stem count)	0.26	0.29	0.52	1.12	0.11	0.26
Shannon Index (by BA)	0.37	0.05	0.09	0.64	0.14	0.27
Simpson Index (by BA)	0.60	0.51	0.98	0.97	0.04	0.09
Average height of the first live branch (m)	8.07	0.92	1.75	6.27	2.68	6.39
STDev. of the height of the first live branch (m)	3.16	0.81	1.62	4.29	2.38	5.16
Average crown horizontal area (m2)	35.14	12.15	22.95	59.58	21.65	46.07
STDev. Crown horizontal area (m2)	25.36	3.03	5.41	33.69	6.96	13.57
Total Crown horizontal area (m2)	1100.99	377.48	737.56	1514.66	254.83	592.78
Standing dead wood volume (m3)	0.14	0.24	0.41	0.46	0.85	1.74
Standing dead wood decay class				0.75	0.12	0.17
Downed dead wood volume (m3)	0.97	0.25	0.49	10.90	18.37	37.81
Downed dead wood decay class	0.71	0.13	0.25	0.70	0.02	0.04
No. saplings per 30x30m	3.33	2.89	5.00	27.75	53.51	108.00
No. native saplings per 30x30m	2.67	2.52	5.00	0.25	0.50	1.00
Percentage of native saplings per 30x30m	53.33	50.33	100.00	25.00	50.00	100.00
No. of sapling species per 30x30m	1.33	1.15	2.00	1.00	0.82	2.00
No. seedlings per 30x30m	18.00	0.00	0.00	272.25	188.12	405.00
No. native seedlings per 30x30m	18.00	0.00	0.00	135.00	146.42	315.00
Percentage of native seedlings per 30x30m	100.00	0.00	0.00	60.19	37.90	86.67
No. of seedling species per 30x30m	1.33	0.58	1.00	5.25	2.99	7.00
Seedlings Shannon Index for native species	0.00	0.00	0.00	0.92	0.74	1.61
No. Vascular Species	3.33	1.15	2.00	5.50	2.89	5.00
No. Bryophyte Species	11.67	5.77	10.00	18.75	2.50	5.00
Percentage of bare soil within 30x30m	3.50	5.63	10.00	51.25	35.21	85.00

5.1.1 Statistical comparison of 2010 and 2012 fieldwork populations

The two fieldwork populations, i.e. those plots enumerated in 2010 and those in 2012, were statistically compared using the Mann-Whitney-Wilcoxon and Kolmogorov-Smirnov two-independent-samples tests. Each of the field metrics were considered as a separate population. The null hypothesis was that the two populations were identical. The hypothesis was tested by applying the two independence tests using R statistical software. The $p < 0.05$ significance level was used, thus if the p value met this criterion the null hypothesis is rejected.

Table 5.9 presents the Mann-Whitney-Wilcoxon and Kolmogorov-Smirnov test results. The majority of field metric populations have p values for both statistical measures above 0.05; therefore the null hypothesis is retained. Seven of the field metric populations were significantly different at $p < 0.05$ for both statistical tests. These metrics were the standard deviation of tree NN spacing, the Simpson index (by stem count), DDW decay class, the number of native saplings, the percentage of native saplings, the number of seedling species and the Shannon index for native seedling species. The majority of these metrics relate to understorey composition.

Five field metrics had significance values of $p < 0.05$ for the Mann-Whitney-Wilcoxon test only. These were the average tree NN spacing, the number of sapling species, the number of seedlings, the percentage of native seedlings, and number of bryophyte species. Only one of the field metrics had significance values of $p < 0.05$ for the Kolmogorov-Smirnov test only, this was the total DDW volume.

Table 5.9 – Statistical comparisons between 2010 and 2012 fieldwork populations

Metric name	Mann-Whitney-Wilcoxon		Kolmogorov-Smirnov	
	W	p	D	p
<i>No. Canopy Layers</i>	225.00	0.68	0.06	1.00
<i>Number of trees per 30x30m</i>	164.00	0.23	0.18	0.88
<i>Number of native trees per 30x30m</i>	173.50	0.35	0.23	0.65
<i>Percentage of native trees</i>	218.00	0.84	0.12	1.00
<i>No. of tree species per 30x30m</i>	203.50	0.87	0.10	1.00
<i>Average tree spacing (m)</i>	275.00	0.04	0.37	0.10
<i>STDev. of tree spacing (m)</i>	302.00	0.00	0.52	0.00
<i>Mean DBH (cm)</i>	275.00	0.09	0.38	0.06
<i>STDev. of DBH (cm)</i>	197.00	0.75	0.22	0.63
<i>Basal area within 30x30m</i>	225.00	0.71	0.18	0.80
<i>Percentage big trees (DBH 40>x<80cm)</i>	260.00	0.20	0.28	0.38
<i>Percentage very big trees (DBH >80cm)</i>	228.00	0.53	0.24	0.61
<i>Percentage big trees (DBH >40cm)</i>	269.00	0.13	0.32	0.24
<i>Loreys mean height (m)</i>	170.00	0.31	0.21	0.68
<i>Shannon index (by stem count)</i>	226.00	0.69	0.21	0.75
<i>Simpson index (by stem count)</i>	113.00	0.01	0.56	0.00
<i>Shannon index (by BA)</i>	219.00	0.82	0.23	0.65
<i>Simpson index (by BA)</i>	163.00	0.23	0.27	0.44
<i>Average height of the first live branch (m)</i>	211.00	0.78	0.20	0.83
<i>STDev. of the height of the first live branch (m)</i>	153.00	0.21	0.27	0.48
<i>Average crown horizontal area (m2)</i>	231.00	0.41	0.29	0.29
<i>STDev. crown horizontal area (m2)</i>	202.00	0.96	0.18	0.82
<i>Total crown horizontal area (m2)</i>	182.00	0.86	0.20	0.75
<i>Standing deadwood volume (m3)</i>	228.00	0.64	0.27	0.43
<i>Standing deadwood decay class</i>	107.00	0.68	0.39	0.24
<i>Downed deadwood volume (m3)</i>	161.00	0.21	0.42	0.04
<i>Downed deadwood decay class</i>	316.00	0.01	0.52	0.01
<i>No. saplings per 30x30m</i>	151.50	0.09	0.47	0.09
<i>No. native saplings per 30x30m</i>	186.50	0.00	0.67	0.00
<i>Percentage of native saplings per 30x30m</i>	187.50	0.00	0.75	0.00
<i>No. of sapling species per 30x30m</i>	159.50	0.04	0.40	0.21
<i>No. seedlings per 30x30m</i>	119.00	0.02	0.40	0.07
<i>No. native seedlings per 30x30m</i>	145.00	0.09	0.42	0.05
<i>Percentage of native seedlings per 30x30m</i>	302.00	0.00	0.41	0.07
<i>No. of seedling species per 30x30m</i>	92.00	0.00	0.56	0.00
<i>Seedlings Shannon index for native species</i>	83.50	0.00	0.61	0.00
<i>No. vascular species</i>	171.50	0.31	0.16	0.95
<i>No. bryophyte species</i>	117.50	0.01	0.38	0.10
<i>Percentage of bare soil within 30x30m</i>	154.00	0.14	0.36	0.14
<i>Sum of DBH differences (cm)</i>	245.00	0.21	0.37	0.44
<i>Index of vertical separation</i>	147.00	0.16	0.28	0.35

[Underlined values indicate the populations are significantly different ($p < 0.05$)]

5.2 Relationships between field metrics

Two statistical processes were implemented to determine if there were any relationships between each of the field metrics (for the 2010 field data), first with a bivariate correlation and secondly with a standard OLS multiple regression.

5.2.1 Bivariate correlation

A bivariate correlation matrix using the spearman's rho was used to investigating correlations between the 37 metrics of the 2010 dataset. This matrix of correlation coefficients and significance (p) values are presented in Table 5.10. Each of the correlation pairs were classified into one of four arbitrary classes based upon the correlation coefficient value (where $p < 0.05$), where:

1. strong ($r \geq 0.8$ or $r \leq -0.8$) – blue;
2. medium ($0.6 \geq r < 0.8$ or $-0.8 > r \geq -0.6$) – green;
3. weak ($0.4 \geq r < 0.6$ or $-0.6 > r \geq -0.4$) – grey;
4. none ($0 \geq r < 0.4$ or $-0.4 > r < 0$) – white.

A total of 8 metric pairs received a 'high' classification, 43 were 'moderate', and 75 were 'low'. It is clear from Table 5.10 that there is a great deal of interrelationship between the field metrics. For example mean DBH has strong correlations with the total number of trees and the average crown horizontal area within the 30x30m plot, moderate correlations with average tree nearest neighbour (NN) spacing and the percentage of big trees, whereas weak correlations were observed for the standard deviation of crown area, standing deadwood (snag) volume and downed deadwood (DDW) volume. Average crown horizontal area medium correlations with the number of trees, average tree NN spacing, percentage very big trees and snag decay class, in addition to weak correlations with the standard deviation of tree spacing, percentage of all big trees, average height of the first live branch, standing deadwood volume and DDW decay class. The Shannon index (from basal area) has a high correlation with the number of saplings present in the plot. Total basal area has a medium correlation with total crown horizontal area within the plot.

Strong correlations were identified between the number of sapling and number of native saplings, and between the number of seedlings and the number of native seedlings. This is probably due to many of the seedlings and saplings encountered in the field being of native species.

5.2.2 OLS multiple regression

For the 2010 field data, a total of 15 field metrics out of the 39 could not be estimated using an OLS multiple regression approach, i.e. there were no significant relationships. Table 5.11 summarises the 24 models which could be produced. Of the metrics used in the significant models, mean DBH was the most frequently used predictor (for estimating the number of trees, percentage of very big trees and horizontal crown dimensions). The second most frequent predictor was average horizontal crown area, which had relationships with mean and standard deviation of DBH, average crown base height and the volume of standing deadwood. A number of the other metrics were used in two models, such as canopy openness, which was related with the number of trees and total crown horizontal area. Canopy openness could not be modelled by any combination of other field metrics, nor could basal area, DDW class, number of canopy layers, number of tree species and by extension the Shannon and Simpson indices of diversity and the number of native trees. The number of native saplings and number of vascular species could also not be estimated.

Table 5.10 – Spearman’s rho bivariate correlation matrix (colours indicate strength of the correlation: blue = strong; green = medium; grey = weak; white = none)

	Canopy openness	Can. Layers	No. trees	No. native trees	Tree species	Av. Tree NN space	StDev tree NN space	Mean DBH	STD DBH	Total BA	%big trees (40-80cm)	% big trees (80+)	%all big trees (40+)	Av. height	SH (stem count)	SI (stem count)	SH (byBA)	SI (byBA)	Av. Ht. Live crown	STD Ht. Live crown	Av. crown area	STD crown area	Total CA	Standing dead vol.	Standing dead decay	DDW vol.	DDW decay	No. saplings	No. Nat. saplings	Sapling species	No. seedlings	No. Nat. seedlings	Seedling species	Seedlings SH	Vascular species	Bryophyte species	% Bare soil		
Canopy openness	1.000	.000	.057	-.152	.355	-.059	-.036	-.191	-.426	-.343	-.057	-.342	-.065	-.167	.333	.073	.056	.281	.326	.341	-.364	-.639	-.383	-.515	-.052	-.314	-.310	.065	-.106	.362	-.170	-.321	.273	.090	-.232	-.477	-.289		
Can. Layers		1.000	.185	.352	.482	-.161	.031	-.326	.225	.213	-.386	.003	-.354	-.446	.594	.575	.112	.325	-.429	.024	-.115	.178	.150	.022	-.525	-.101	-.004	-.222	-.352	.328	-.228	-.232	-.072	-.315	-.109	-.028	.320		
No. trees			1.000	.325	.243	-.702	-.604	-.889	-.388	.245	-.581	-.617	-.647	-.242	.102	-.014	.210	.227	.357	.392	-.782	-.502	.348	-.378	-.467	-.330	-.160	-.062	-.145	.070	-.208	-.193	-.276	-.332	-.271	-.084	.039		
No. native trees				1.000	.414	-.422	-.014	-.277	.247	.355	-.349	-.010	-.347	-.091	.233	.133	.064	.226	-.194	.352	-.088	.135	.345	.103	-.119	.204	.128	.032	-.106	.150	-.131	-.061	-.180	-.222	-.278	.050	.225		
No. Tree species					1.000	-.304	.193	-.348	.260	.135	-.167	-.158	-.196	-.214	.901	.612	.327	.363	-.344	.376	-.205	-.117	.133	-.054	-.107	-.136	-.096	.073	-.061	.562	.284	.243	.460	.223	-.105	-.082	.321		
Av. Tree NN space						1.000	.153	.671	-.051	-.284	.682	.146	.724	.292	-.192	-.142	-.179	-.158	.051	-.457	.635	.318	-.175	.083	.263	.084	.318	-.028	.119	-.246	-.046	-.047	.070	.269	.082	-.020	-.303		
StDev tree NN space							1.000	.430	.666	-.156	.008	.684	.093	-.129	.164	.209	-.421	-.101	-.639	-.358	.503	.552	-.238	.381	.454	.066	.146	.576	.569	.454	.220	.247	.400	.217	.271	.256	.234		
Mean DBH								1.000	.406	.092	.656	.700	.731	.342	-.256	-.109	-.209	-.276	-.250	-.316	.845	.564	-.058	.491	.504	.468	.236	-.189	-.005	-.295	.164	.199	.147	.225	.328	.201	.061		
StDev DBH									1.000	.357	.080	.727	.151	-.062	.224	.324	-.152	.049	-.768	-.192	.412	.669	.025	.771	.370	.301	.242	.028	.092	.222	.425	.513	.406	.178	.275	.476	.557		
Total BA										1.000	-.103	.381	-.051	.035	.086	.105	.075	-.036	-.086	.205	.065	.247	.621	.309	-.327	.388	.094	-.300	-.202	.183	-.017	.075	-.093	-.155	.235	.213	.435		
%big trees (40-80cm)											1.000	.087	.980	.636	-.150	-.284	-.013	-.389	.220	.046	.458	.057	-.253	.134	.154	.312	.129	-.215	-.129	-.450	.227	.214	.270	.473	.330	.124	-.023		
% Big trees (80+)												1.000	.206	.022	-.072	.135	-.319	-.203	-.531	-.326	.653	.701	.047	.644	.205	.412	.229	.132	.233	.132	.143	.205	.045	-.098	.465	.247	.481		
% all big trees (40+)													1.000	.650	-.169	-.285	-.125	-.380	.176	.012	.531	.138	-.220	.168	.226	.346	.231	-.104	-.005	-.356	.203	.191	.274	.455	.342	.155	.023		
Av. height														1.000	-.246	-.434	-.188	-.441	.567	.340	.142	-.064	-.062	.069	.038	.475	.047	.304	.317	-.010	.295	.325	.107	.426	.371	.042	.009		
SH (stem count)															1.000	.793	.502	.437	-.407	.211	-.075	-.034	.152	-.025	-.155	-.140	-.082	-.141	-.225	.431	.266	.190	.356	.125	-.151	-.138	.366		
SI (stem count)																1.000	.426	.584	-.628	-.097	.139	.211	.228	.264	.075	-.057	.057	-.312	-.356	.295	.349	.293	.286	-.036	-.183	-.083	.283		
SH (byBA)																	1.000	.364	-.137	.123	-.138	-.215	.262	-.066	-.208	-.111	-.117	-.876	-.734	-.439	.189	.111	-.039	.041	-.327	-.049	.092		
SI (byBA)																																							
Av. Ht. Live crown																			1.000	.451	-.471	-.666	-.113	-.585	-.511	-.108	-.191	.333	.282	-.013	-.281	-.312	-.304	-.013	.124	-.363	-.271		
STD. ht. Live crown																				1.000	-.490	-.569	.012	-.365	-.390	.133	-.250	.268	.149	.404	.131	.069	.103	.121	.087	-.119	.248		
Av. crown area																					1.000	.736	.209	.447	.645	.423	.456	-.198	-.110	-.352	.085	.103	.052	.084	.152	.285	.059		
STD crown area																						1.000	.300	.648	.498	.182	.267	-.097	-.005	-.101	.062	.207	-.008	-.043	.115	.418	.274		
Total CA																							1.000	.061	.148	.192	.291	-.456	-.422	-.179	-.116	-.042	-.285	-.224	-.281	.312	.170		
standing dead vol.																									1.000	.260	.413	.214	-.298	-.347	-.146	.415	.537	.207	.026	.249	.399	.303	
standing dead decay																											1.000	.049	.372	.359	.359	.029	.396	.411	.493	.444	-.400	.378	-.244
DDW vol.																											1.000	.141	.120	.124	.266	.310	.326	.041	.131	.301	.125	.005	
DDW decay																												1.000	.104	.069	-.344	.007	-.059	-.024	-.056	-.084	.411	.150	
No. saplings																													1.000	.930	.668	.196	.232	.188	.184	.291	-.146	.070	
No. Nat. saplings																														1.000	.631	.268	.304	.149	.289	.311	-.174	.063	
Sapling species																															1.000	.450	.451	.479	.334	.229	-.359	.068	
No. seedlings																																1.000	.961	.698	.655	.219	.117	.104	
No. Nat. Seedlings																																	1.000	.638	.625	.221	.183	.130	
Seedling species																																		1.000	.827	.138	.200	-.105	
Seedlings SH																																			1.000	.122	.208	-.289	
Vascular spec.																																				1.000	.023	.388	
Bryophyte spec.																																					1.000	.186	
% Bare soil																																						1.000	

Table 5.11 – OLS regression model relationships between field metrics. (All models were significant at $p < 0.05$)

Metric	R ²	Variables in regression equation		
No. trees per 30x30m plot	0.826	MeanDBH	Canopy openness	-
Average tree spacing (m)	0.827	(iii) Percentage big trees	StDev. crown base height	DDW volume
StDev. tree spacing (m)	0.677	StDev. horizontal crown area	Total horizontal crown area	-
MeanDBH (cm)	0.926	Avg_crown	(iii) Percentage big trees	Average crown base height
StDev. DBH (cm)	0.706	StDev. crown base height	Average horizontal crown area	-
(i) Percentage big trees (DBH 40>x<80cm)	0.625	Snag volume	-	-
(ii) Percentage very big trees (DBH >80cm)	0.774	STDev. DBH	-	-
(iii) Percentage big trees (DBH >40cm)	0.749	Mean DBH	StDev.DBH	-
Average crown base height (m)	0.587	StDev. Tree spacing	-	-
StDev. crown base height (m)	0.284	StDev. horizontal crown area	-	-
Average horizontal crown area (m ²)	0.789	Mean DBH	-	-
StDev. horizontal crown area (m ²)	0.914	Mean DBH	-	-
Total horizontal crown area (m ²)	0.404	Canopy openness	Average tree spacing	-
Standing deadwood volume (m ³)	0.485	Average horizontal crown area	-	-
Standing deadwood class	0.821	No seedlings per 30x30m plot	-	-
DDW (m ³)	0.895	Mean DBH	-	-
No. Sapling species	0.992	StDev. tree spacing	No. tree species	-
No. seedlings per 30x30m plot	0.403	Snag class	-	-
No. Native seedlings per 30x30m plot	0.986	Snag class	(ii) Percentage very big trees	-
Percentage native seedlings per 30x30m plot	0.606	DDWclass		-
No. seedlings species	0.446	No seedlings per 30x30m plot	Total horizontal crown area	-
Native seedlingsShannon index	0.966	No. seedlings per 30x30m plot	No. Tree species	-
No. Moss species	0.635	Shannon index (stem count)	-	-
Percentage cover of bare soil	0.678	No. seedlings species	-	-
Field metrics for which no model could be produced: Canopy openness (%); No. canopy layers; No. native trees per 30x30m plot; Percentage of native trees per 30x30m plot; No. Tree species; Total BA per 30x30m plot (m ²); Average tree height (m), Shannon index (stem count); Simpson index (stem count); Shannon index (BA); Simpson index (BA); DDWclass; No. saplings per 30x30m plot; No. Native saplings per 30x30m plot; Percentage native saplings per 30x30m plot; and No. vascular species.				

5.3 Condition index results for field plots

The compositional and structural indices, as identified in section 4.2.3, were calculated for both the 2010 and 2012 fieldwork datasets. A full list of the indices produced for 2010 and 2012 data are presented in Tables 5.12 and 5.13 respectively. The compositional indices, the Shannon (SH) and the Simpson (SI) varied a great deal between field plots. For example in plot 6, there was only a single tree species, oak, which resulted in SH and SI scores of 0. The SH and SI calculated for the proportions of stem numbers of a species and the total basal area of that species varied significantly. For plot 7, the SH and SI of basal area were less than for stem count, because the relative proportions differed between stem count and total basal area for the same species. For example, Douglas fir proportions change from 52% (stem count) to 74% (basal area). For plot 10, SH and SI behaved differently, but the cause was the same; as beech was the most numerous species (55%), but accounted for a small basal area (17%).

The aggregation index varied from 0.5 to 1.65 amongst the 41 plots, indicating a tendency towards clustering for the former, and a tendency towards a regular hexagonal distribution of trees for the latter. It should be noted that the 2012 dataset contained five plots with an index value less than 1, compared with just one such plot in the 2010 data. For example plot number 26, which is a mix of deciduous species of various sizes located in semi-ancient woodland, had an index score of 0.52. In contrast, plot 31, which is located within a single species coniferous plantation where many if not all the trees were planted at the same time, had an index score of 1.45.

The DBH or diameter differentiation index (TM) illustrates the differentiation between neighbouring trees DBH values, where values for 2010 ranged from 0.18 to 0.99 indicating stands with smaller diameter variability and high variability, respectively.

The vertical evenness (VE) index characterises the vertical distribution of horizontal cover within a plot within four height layers. These height layers correspond to percentages relative to the maximum tree height recorded within the plot, the divisions being: (i) 0-19%; (ii) 20-49%; (iii) 50-79%; and (iv) 80-100%. Within the dataset VE ranged from 0.28 to 1.04. Low values characterise single storied stands while index values closer to 1 would be indicative of trees which were evenly vertically distributed. For example, plot number 27, is within deciduous woodland and has approximately three canopy layers present, with a height range of 7m to 27m, with significant vegetative cover within each of these height layers, giving it a

VE index of 0.68. Plot number 31 by contrast had a VE index score of 0.28. Plot number 31 is a single species dominated coniferous plantation, with a limited height range of 27 to 32 m and while there is a great deal of vegetation within the upper strata, there is very little at lower levels.

The complexity index (H) represents the first of the composite indices, which takes into account number of tree stems, the number of tree species, dominant height and basal area. For this study, the index values range from 6.5 to 567.4. For example, plot number 39 exhibits a small index value, 14.8, due to a relatively low stem density, being composed of only one tree species and a relatively low basal area value (2.3m^2 per $30\text{x}30\text{m}$), but high average height (33m). In contrast plot 30 had the highest value of 567.4, and was a plot characterised by five different tree species at high density (90 stems within $30\text{x}30\text{m}$), and high basal area value (5.2m^2 per $30\text{x}30\text{m}$), although with a relatively small dominant height (19m).

The second composite index was the complex stand diversity index (CSDI). This is constructed from four separate indices, which include measures of species diversity, DBH differentiation, nearest neighbour proximity, and crown dimensions. Values derived from field data ranged from 2.3 to 12.8, where values less than 4 are considered indicative of homogeneous woodland, and values above 9 indicate a very heterogeneous structure. For example, plot 12 had a low index value of 3.29 due to the plot being dominated by a single species, Corsican pine, with little DBH differentiation, and tree spacing and crown dimensions differing only by 3-5 metres. An example of a plot with a high index value (12.78) is number 24, where the plot contains three tree species, DBH values can vary by ~40cm, neighbouring tree distances can vary up to 7m, and crown dimensions up to 10m.

Table 5.12 – Indices derived from 2010 field data

Plot no.	Aggregation Index (R)	Diameter differentiation Index (TM)	Vertical Evenness (VE)	Complexity Index (HC)	Complex Stand diversity index (CSDI)	Shannon Index (Tree Stem count)	Shannon Index (Basal Area)	Evenness Index (E)	Simpson Index (Tree Stem count)	Simpson Index (Basal Area)	Class
Plot 1	1.08	0.31	0.47	52.45	6.16	0.16	0.42	0.16	0.07	0.42	Mixed
Plot 2	1.47	0.19	0.67	41.95	4.20	0.24	0.35	0.24	0.12	0.20	Mixed
Plot 3	1.36	0.22	1.01	34.67	6.84	1.19	1.02	0.60	1.00	0.59	Conif.
Plot 4	1.06	0.52	1.04	27.20	6.55	0.99	0.24	0.63	0.67	0.95	Decid.
Plot 5	1.31	0.26	0.50	77.65	5.00	0.88	0.76	0.56	0.57	1.00	Conif.
Plot 6	1.49	0.25	0.41	6.57	2.35	0.00	0.00	0.00	0.00	0.00	Decid.
Plot 7	1.42	0.35	1.00	44.33	6.28	1.15	0.85	0.57	0.67	0.74	Conif.
Plot 8	1.16	0.36	0.99	109.17	6.61	1.13	1.02	0.57	0.65	0.76	Conif.
Plot 9	1.17	0.32	1.04	27.23	6.44	0.93	0.75	0.59	0.64	0.88	Decid.
Plot 10	0.79	0.42	0.90	116.92	6.46	0.85	0.57	0.53	0.56	0.40	Decid.
Plot 11	1.05	0.45	0.88	177.13	8.00	1.14	0.47	0.49	0.75	0.89	Decid.
Plot 12	1.58	0.19	0.56	36.41	3.29	0.33	0.49	0.33	0.00	0.47	Conif.
Plot 13	1.19	0.42	1.02	52.76	5.83	0.83	0.48	0.52	0.54	0.37	Decid.
Plot 14	1.13	1.00	1.00	76.95	7.11	0.94	0.30	0.47	0.52	0.19	Conif.
Plot 15	1.39	0.34	0.87	145.16	7.15	1.07	1.14	0.53	0.55	0.87	Decid.
Plot 16	1.30	0.26	0.84	144.88	7.96	1.16	0.67	0.45	0.28	0.39	Conif.
Plot 17	1.65	0.23	0.54	92.26	5.01	0.59	0.35	0.37	0.34	0.27	Conif.
Plot 18	1.57	0.20	0.93	73.07	5.56	0.84	0.43	0.53	0.51	0.72	Conif.
Plot 19	1.02	0.36	0.93	113.20	7.28	1.13	0.33	0.49	0.59	1.18	Mixed
Plot 20	1.19	0.40	0.97	97.71	7.90	1.47	0.54	0.63	0.80	0.42	Decid.
Plot 21	1.39	0.32	0.97	173.24	7.67	1.07	0.70	0.68	0.66	0.53	Conif.

Table 5.13– Indices derived from 2012 field data

Indices	Aggregation Index (-R)	Diameter differentiation index (-TM)	Vertical Evenness (VE)	Complexity Index (HC)	Complex Stand diversity index (CSDI)	Shannon Index (Tree Stem count)	Shannon Index (Basal Area)	Evenness Index (E)	Simpson Index (Tree Stem count)	Simpson Index (Basal Area)	Class
PLOT 22	-	0.19	0.32	45.93	-	1.39	0.90	0.60	0.75	0.77	Decid.
PLOT 23	1.34	0.22	0.38	67.46	11.50	0.65	0.39	0.65	0.46	0.41	Conif.
PLOT 24	0.97	0.52	-	88.47	12.78	0.46	0.86	0.29	0.24	0.86	Decid.
PLOT 25	0.70	0.26	0.48	59.63	9.92	0.98	0.81	0.62	0.61	0.75	Decid.
PLOT 26	0.52	0.25	0.44	64.69	10.22	1.15	0.97	0.50	0.63	0.80	Decid.
PLOT 27	0.72	0.35	0.69	109.91	11.47	0.96	0.77	0.60	0.59	0.71	Decid.
PLOT 28	1.22	0.36	0.37	109.48	6.88	0.74	0.75	0.47	0.48	0.66	Conif.
PLOT 29	0.83	0.32	0.54	38.56	10.46	0.54	0.68	0.54	0.37	0.73	Decid.
PLOT 30	1.17	0.42	-	567.43	-	0.98	1.29	0.42	0.57	0.84	mixed
PLOT 31	1.45	0.45	0.29	28.22	4.27	0.00	0.00	0.00	0.00	0.00	Conif.
PLOT 32	1.34	0.19	0.40	100.28	9.22	1.35	1.24	0.67	0.77	0.93	mixed
PLOT 33	1.41	0.42	0.39	137.89	9.51	0.74	0.30	0.37	0.38	0.20	Conif.
PLOT 34	1.24	1.00	0.56	106.47	12.39	1.43	1.15	0.62	0.76	1.01	Conif.
PLOT 35	1.17	0.34	0.48	194.75	10.15	1.02	0.67	0.51	0.60	0.51	Conif.
PLOT 36	1.42	0.26	0.33	92.20	7.42	0.82	0.58	0.41	0.46	0.41	mixed
PLOT 37	1.14	0.23	0.50	45.53	8.38	0.78	0.50	0.49	0.50	1.60	Decid.
PLOT 38	1.20	0.20	-	120.84	8.78	0.90	1.09	0.57	0.54	0.97	mixed
PLOT 39	1.38	0.36	0.52	14.84	5.76	0.00	0.00	0.00	0.00	0.00	Conif.
PLOT 40	1.21	0.40	0.50	31.78	8.11	0.87	0.70	0.55	0.54	2.89	Decid.
PLOT 41	1.09	0.79	0.58	78.52	10.64	1.41	0.99	0.61	0.50	0.97	mixed

The third composite index (score-based), was computed from a total of 17 metric scores (Cantarello and Newton, 2008). It should be noted that a number of metrics had to be modified to account for the requirements of the scoring system by extrapolating attributes relating to counts or totals for areas to one hectare. The premise of the approach is based on plot values exceeding targets of ‘ideal’ plot conditions, these targets are listed in both Tables 5.14 and 5.15 along with plot values, for 2010 and 2012 fieldwork respectively. If a target was exceeded the field plot was scored as one, otherwise it received a zero. A plot meeting or exceeding all targets will get a maximum score of seventeen.

For the 41 field plots there is a range of values from 3 to a maximum of 13. Generally, coniferous plots had lower index scores than deciduous plots. The lowest score of 3 was determined for plot number 23, which is a coniferous stand dominated by hybrid larch. The three targets exceeded were the number of trees within the plot, basal area and standing deadwood decay class. Plot number 15, had the lowest value for a deciduous plot, with a score of 6, this plot however showed evidence of active management by forestry personnel, which may explain this relatively low score. The six targets exceeded were the number of trees, basal area, standard deviation of DBH, mean height, and standing and downed deadwood decay classes.

The plot with the highest index value, 13, was plot number 27. This plot was located within isolated old growth deciduous woodland. All targets were exceeded apart from average DBH, number of seedlings and number of native seedlings. Plot number 35 had an index score of 10, this was the highest value for coniferous woodland. The plot was located within plantation woodland, containing a mix of coast redwood (*sequoia sempervirens*) and Douglas fir. The seven targets the plot failed on were species diversity, total number of saplings (and natives), standing deadwood volume, number of seedlings (and natives), and finally the number of ground vegetation species.

Table 5.14 – Score-based condition assessment for 2010

Indicator	Target	Sites																			
1 No. of trees (ha-1)	222	333.33	344.44	211.11	122.22	333.33	77.78	188.89	400.00	122.22	222.22	422.22	222.22	200.00	433.33	488.89	266.67	288.89	233.33	400.00	211.11
2 Shannon index for native trees (H)	0.87	0.47	0.00	0.86	0.99	0.41	0.00	0.68	0.64	0.33	0.85	1.07	0.00	0.83	0.26	0.53	0.95	0.59	0.38	0.86	1.47
3 Basal Area (m2 ha-1)	23	28.79	33.18	23.28	42.42	32.39	25.13	26.25	40.37	26.41	57.20	33.55	33.14	40.82	30.29	45.46	36.13	45.34	39.43	30.22	54.00
4 Mean diameter (cm)	32	30.66	34.28	35.85	57.21	34.37	61.84	40.26	33.19	48.21	49.85	30.42	42.78	45.61	23.90	30.85	39.46	43.74	44.62	27.30	43.69
5 Standard dev. of diameters (cm)	14	12.86	7.30	11.20	33.89	7.62	18.39	12.57	13.73	21.68	28.88	15.72	8.50	23.44	18.09	15.40	13.24	9.40	12.96	5129.09	37.72
6 Percentage of big trees	7	0.00	0.00	0.00	36.36	0.00	14.29	0.00	0.00	0.00	15.00	0.00	0.00	11.11	0.00	0.00	0.00	0.00	0.00	15.79	0.00
7 Mean height (m)	17	18.22	18.35	17.64	17.49	23.97	33.60	20.27	16.90	28.12	30.66	17.86	24.72	21.54	22.93	17.51	25.07	23.48	26.48	20.46	21.51
8 No. of native saplings (ha-1)	91	22.22	0.00	0.00	0.00	0.00	0.00	0.00	11.11	11.11	0.00	30.00	44.44	0.00	27.78	0.00	0.00	0.00	0.00	0.00	0.00
9 No. of native saplings (ha-1)	91	22.22	0.00	0.00	0.00	0.00	0.00	0.00	11.11	11.11	0.00	30.00	44.44	0.00	24.44	0.00	0.00	0.00	0.00	0.00	0.00
10 Vol. downed dead wood (m3 ha-1)	26	10.10	13.91	1.80	9.15	8.67	10.28	16.37	3.30	97.60	162.40	5.42	1.02	1202.26	5.38	3.70	84.56	104.73	11.14	8.44	101.41
11 Downed dead wood decay class	0.5	1.76	2.47	2.71	2.72	2.53	2.86	2.68	2.44	2.57	2.26	2.65	2.50	2.36	2.07	2.80	2.07	2.77	2.07	2.44	2.25
12 Volume of snag (m3 ha-1)	16	0.00	0.00	0.00	47.16	0.80	47.34	0.00	3.14	112.42	76.90	1.82	0.00	40.55	0.00	0.33	0.00	0.00	6.30	4.53	49.40
13 Snag Decay class	0.4	0.00	0.00	0.00	2.33	2.20	2.33	0.00	2.00	2.75	2.50	2.88	0.00	450.61	0.00	2.00	0.00	0.00	2.00	2.00	1.27
14 No. of total seedlings (ha-1)	63219	200.00	200.00	200.00	200.00	300.00	500.00	9600.00	600.00	3700.00	3800.00	44500.00	200.00	1800.00	2400.00	200.00	300.00	100.00	2200.00	100.00	200.00
15 No. of native seedlings (ha-1)	63219	200.00	200.00	200.00	200.00	300.00	500.00	1600.00	500.00	3700.00	3800.00	44500.00	200.00	1800.00	2100.00	200.00	300.00	100.00	2200.00	100.00	200.00
16 Shannon index for native seedlings (H)	0.89	0.00	0.00	0.00	0.00	0.00	0.00	0.66	0.00	0.86	0.39	0.64	0.69	0.45	1.27	0.00	0.64	0.00	0.47	0.00	0.20
17 No. of ground veg. species	33	2.00	5.00	3.00	4.00	3.00	6.00	4.00	4.00	2.00	6.00	2.00	4.00	7.00	7.00	3.00	3.00	4.00	4.00	4.00	6.00
Class	Mixed	Mixed	Conf.	Decid.	Conf.	Decid.	Conf.	Decid.	Conf.	Decid.	Decid.	Decid.	Conf.	Decid.	Conf.	Decid.	Conf.	Conf.	Conf.	Mixed	Decid.
Indicator		Sites																			
1 No. of trees (ha-1)	1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2 Shannon index for native trees (H)	1	1	0	0	1	0	0	0	1	0	1	1	1	0	1	1	1	1	1	1	0
3 Basal Area (m2 ha-1)	1	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	1
4 Mean diameter (cm)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5 Standard dev. of diameters (cm)	0	0	0	0	1	0	1	0	0	1	1	0	1	0	1	0	1	0	0	1	0
6 Percentage of big trees	0	0	0	0	1	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	1
7 Mean height (m)	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
8 No. of total saplings (ha-1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
9 No. of native saplings (ha-1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
10 Vol. downed dead wood (m3 ha-1)	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	1	1	1	0	1
11 Downed dead wood decay class	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12 Volume of snag (m3 ha-1)	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0	0	0	0	1
13 Snag Decay class	0	0	0	0	1	0	1	0	0	1	1	0	0	1	0	0	0	0	0	0	1
14 No. of total seedlings (ha-1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15 No. of native seedlings (ha-1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16 Shannon index for native seedlings (H)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17 No. of ground veg. species	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total scores		4	5	4	9	6	8	4	5	9	10	7	5	9	8	6	7	6	6	6	10

Table 5.15 – Score-based condition assessment for 2012

Indicator	Target	Sites																			
		22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41
1. No. of trees (ha-1)	222	222.22	788.89	588.89	223.33	188.89	366.67	388.89	333.33	1000.00	266.67	222.22	311.11	311.11	366.67	277.78	265.56	400.00	177.78	188.89	266.67
2. Shannon index for native trees (ii)	0.87	1.39	0.00	0.32	0.98	0.99	0.96	0.29	0.54	0.61	0.00	1.07	0.00	0.50	0.56	0.67	0.78	0.89	0.00	0.58	0.95
3. Basal Area (m ² ha-1)	23	32.23	25.38	26.13	38.50	32.54	40.31	40.19	32.33	58.57	36.78	44.00	39.07	29.78	55.60	33.02	33.15	341.29	25.20	22.94	31.00
4. Mean diameter (cm)	32	35.56	18.72	19.52	38.07	40.83	22.76	35.43	30.23	22.88	41.35	47.86	38.54	31.44	37.95	37.23	37.82	30.46	39.61	34.42	33.46
5. Standard dev. of diameters (cm)	14	20.82	7.37	12.50	22.13	20.69	23.96	7.87	17.05	13.80	6.91	15.58	10.85	15.45	22.48	11.53	14.95	12.47	15.87	19.59	55.67
6. Percentage of big trees	7	35.00	0.00	0.68	38.10	58.82	27.27	25.71	23.33	10.00	58.33	70.00	46.43	32.14	27.27	52.00	43.48	19.44	56.25	23.53	100.00
7. Mean Height (m)	17	12.83	16.85	19.17	22.13	21.05	24.79	23.35	17.89	19.38	28.78	25.64	28.36	22.99	23.88	25.13	17.91	21.40	33.12	24.45	16.99
8. No. of total saplings (ha-1)	91	0.00	0.00	0.00	0.00	0.00	422.22	88.89	22.22	166.67	11.11	22.22	0.00	1200.00	44.44	11.11	77.78	0.00	0.00	111.11	11.11
9. No. of native saplings (ha-1)	91	0.00	0.00	0.00	0.00	0.00	422.22	22.22	22.22	44.44	11.11	0.00	0.00	0.00	22.22	11.11	77.78	0.00	0.00	111.11	11.11
10. Vol. downed dead wood (m ³ ha-1)	26	91.84	10.65	11.29	21.68	1920.27	103.52	23.67	20.15	17.95	15.63	6.88	7.40	426.94	29.62	15.84	24.56	19.09	61.15	43.88	31.36
11. Downed dead wood decay class	0.5	0.78	0.49	0.95	0.70	0.77	0.85	0.71	0.69	0.50	0.48	0.71	0.52	0.73	0.65	0.71	0.73	0.89	0.66	0.71	0.69
12. Volume of snag (m ³ ha-1)	16	5.61	1.14	0.99	1.31	15.14	37.95	0.08	0.49	22.11	0.00	19.34	0.24	1.19	1.24	0.13	0.00	0.00	0.82	0.00	0.00
13. Snag Decay class	0.4	0.71	0.67	0.78	0.78	0.67	0.59	0.33	1.00	0.55	0.00	0.67	1.00	0.83	0.73	1.00			0.67		
14. No. of total seedlings (ha-1)	63219	150.00	0.00	50.00	7800.00	8200.00	1040.00	1300.00	50.00	100.00	120.00	2300.00	4488.89	4500.00	8100.00	5800.00	4500.00	40.00	1600.00	7700.00	4900.00
15. No. of native seedlings (ha-1)	63219	150.00	0.00	50.00	7800.00	8000.00	1040.00	700.00	50.00	100.00	100.00	1100.00	44.44	600.00	600.00	5800.00	4500.00	40.00	1600.00	1700.00	3900.00
16. Shannon index for native seedlings (ii)	0.89	0.00	0.00	0.67	1.00	0.35	1.12	0.88	1.33	0.35	1.58	1.41	0.69	1.33	1.05	1.15	0.73	0.00	0.94	1.04	1.79
17. No. of ground veg. species	33	3.00	3.00	4.00	3.00	5.00	3.00	5.00	4.00	6.00	5.00	4.00	2.00	8.00	4.00	6.00	4.00	3.00	8.00	4.00	8.00
Indicator		Decid.	Conf.	Decid.	Decid.	Decid.	Decid.	Conf.	Decid.	Mixed	Conf.	Mixed	Conf.	Conf.	Mixed	Decid.	Decid.	Mixed	Conf.	Decid.	Mixed
1. No. of trees (ha-1)	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1
2. Shannon index for native trees (ii)	1	0	0	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1
3. Basal Area (m ² ha-1)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
4. Mean diameter (cm)	1	0	0	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	0	1	1
5. Standard dev. of diameters (cm)	1	0	0	1	1	1	1	0	1	0	0	1	0	1	1	1	1	0	1	1	1
6. Percentage of big trees	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7. Mean Height (m)	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8. No. of total saplings (ha-1)	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	1	0
9. No. of native saplings (ha-1)	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10. Vol. downed dead wood (m ³ ha-1)	1	0	0	0	0	1	1	1	0	0	0	0	0	1	1	1	0	0	1	1	1
11. Downed dead wood decay class	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12. Volume of snag (m ³ ha-1)	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0
13. Snag Decay class	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	0	0
14. No. of total seedlings (ha-1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15. No. of native seedlings (ha-1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16. Shannon index for native seedlings (ii)	0	0	0	0	1	0	1	0	1	0	1	1	0	1	1	1	1	0	1	1	1
17. No. of ground veg. species	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	9	3	5	10	9	13	6	8	8	8	6	11	7	10	10	8	7	5	9	8	10

5.4 Summary of field plot data results

Field data capture was carried out in the summers of 2010 and 2012 covering a variety of deciduous, coniferous and mixed stands. There are a number of trends evident in the field data which highlight the differences between the woodland types. Deciduous plots generally have a smaller stem count, larger and more variable tree sizes, lower heights to the first live branch, and larger horizontal crown areas in contrast to coniferous plots. Fewer trees within a plot generally indicate presence of larger tree sizes. The volume of standing deadwood encountered within field plots is generally low, although decay class is more variable in coniferous plots, and consistently high in deciduous plots. Higher DDW volume is encountered in deciduous plots, showing similar variability in decay class. Regeneration levels of saplings are low across all plots. Seedling populations were very variable across all plots, although those of deciduous plots are generally higher.

Non-parametric two-sample population tests were employed in order to determine if the 2010 and 2012 fieldwork populations were similar or dissimilar. For 31 of a total of 41 field metrics both tests indicated there was no significant difference between populations. For seven of the remaining metrics both tests indicated the populations were significantly different. These metrics were the standard deviation of DBH value, Shannon index (by stem count), DDW decay class, the number of native saplings, the percentage of native saplings, the number of seedling species, and the native seedlings Shannon index. Six further metrics were highlighted as significant using one test, but not the other. These 13 field metrics are typically related to understorey composition.

A bivariate correlation matrix was constructed where all field metrics were assessed against one another. A high number of significant ($p < 0.05$) correlations were present with correlation coefficient values above 0.4 (or below -0.4). Eight were identified with strong, 39 with medium, and 33 with weak correlations. The metrics which were correlated with the most other metrics were mean DBH and average crown horizontal area. The application of OLS regression between field metrics found that the most critical field measurements were mean DBH, crown horizontal area and canopy openness, each of these being the most frequent predictors of other metrics. 15 of a total of 39 field metrics could not be modelled using the other field metrics in OLS regression.

A total of 11 condition indices were calculated for each plot, all of which showed a great deal of variability. General trends were evident between coniferous and deciduous plot types however, for example the Clark-Evans aggregation index values indicated a regular distribution of tree stems for coniferous plots but a random distribution for deciduous ones.

Chapter 6 – Forest assessment using airborne hyperspectral data

This chapter presents the results of the individual tree crown (ITC) object-based classification of tree species using hyperspectral data. It also reports the regression models produced for estimating the required field plot-level metrics for assessing condition. The regression models used both summary statistics generated from vegetation indices (VI) and ITC classification information.

6.1 Object-based tree species classification

A LiDAR-derived CHM was used in the eCognition image analysis software to automatically delineate ITC objects and crown clusters using a region-growing algorithm based on height maxima. Classification of the ITC objects was then performed using leaf-on and leaf-off Eagle hyperspectral data from 2010. The hyperspectral dataset comprised six MNF transformed bands for leaf-on and six more for leaf-off data. Classification of ITC objects was implemented through a hierarchical method, utilising fuzzy-membership functions producing a total of 26 tree species classes. These included a number of classes specifically for young tree species types, which exhibited a height less than 15m, and often a different signature in the MNF transformed hyperspectral data.

Three large-scale subsets of the entire classified map for the study site are presented in Figures 6.1-6.3. A smaller and more detailed example subset is presented in Figure 6.4. As illustrated in Figure 6.5, there were a number of examples of over-segmentation of tree crown structures caused by the automated ITC delineation approach. This over-segmentation was more prevalent in deciduous areas.

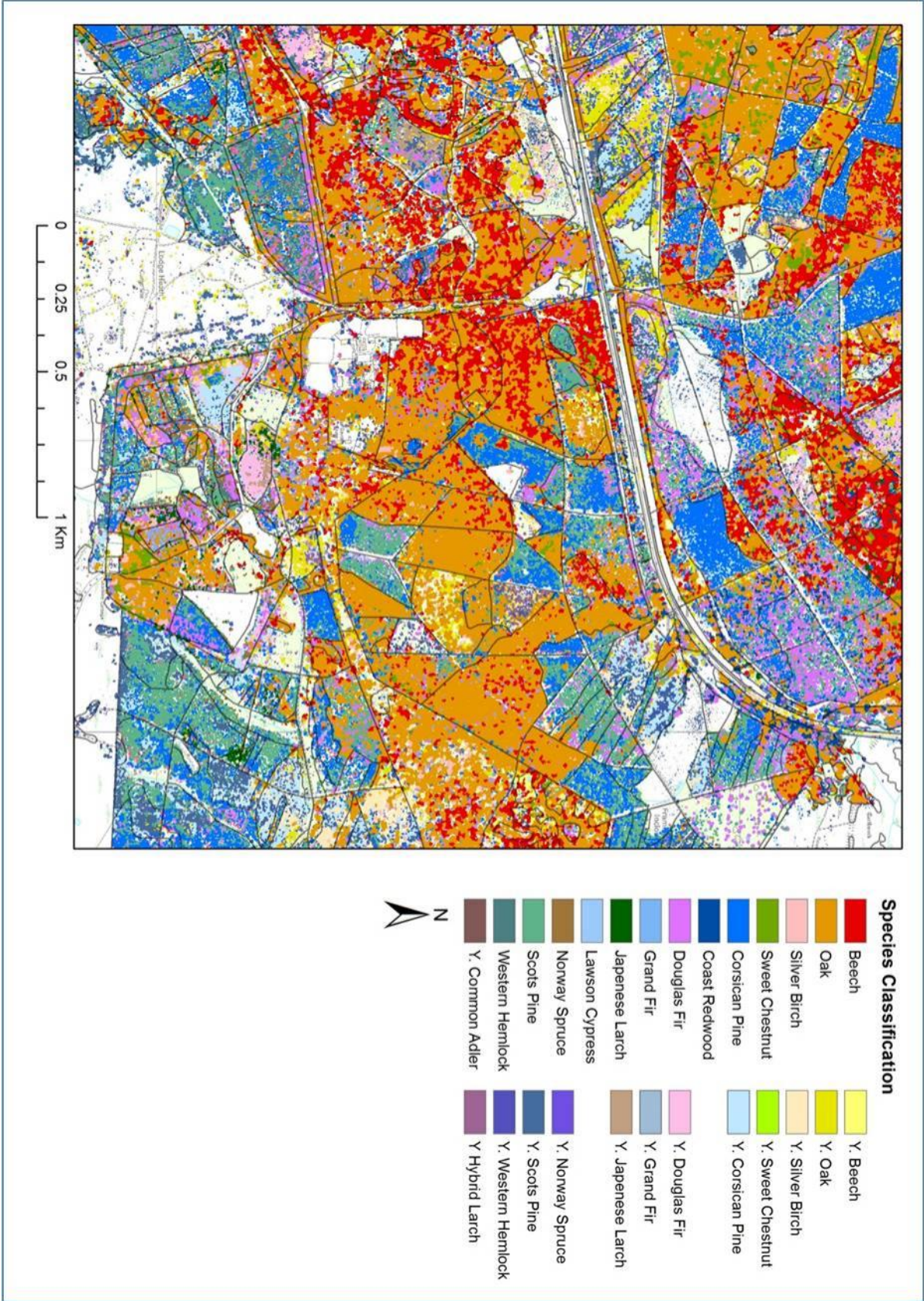


Figure 6.1 – Classification map part 1: Frame-Heath, Lady-Cross, Stockley and New Copse inclosures (Note that “Y.<class_name>” denotes ‘young’, i.e. lower tree height and/or different hyperspectral signature than mature trees)

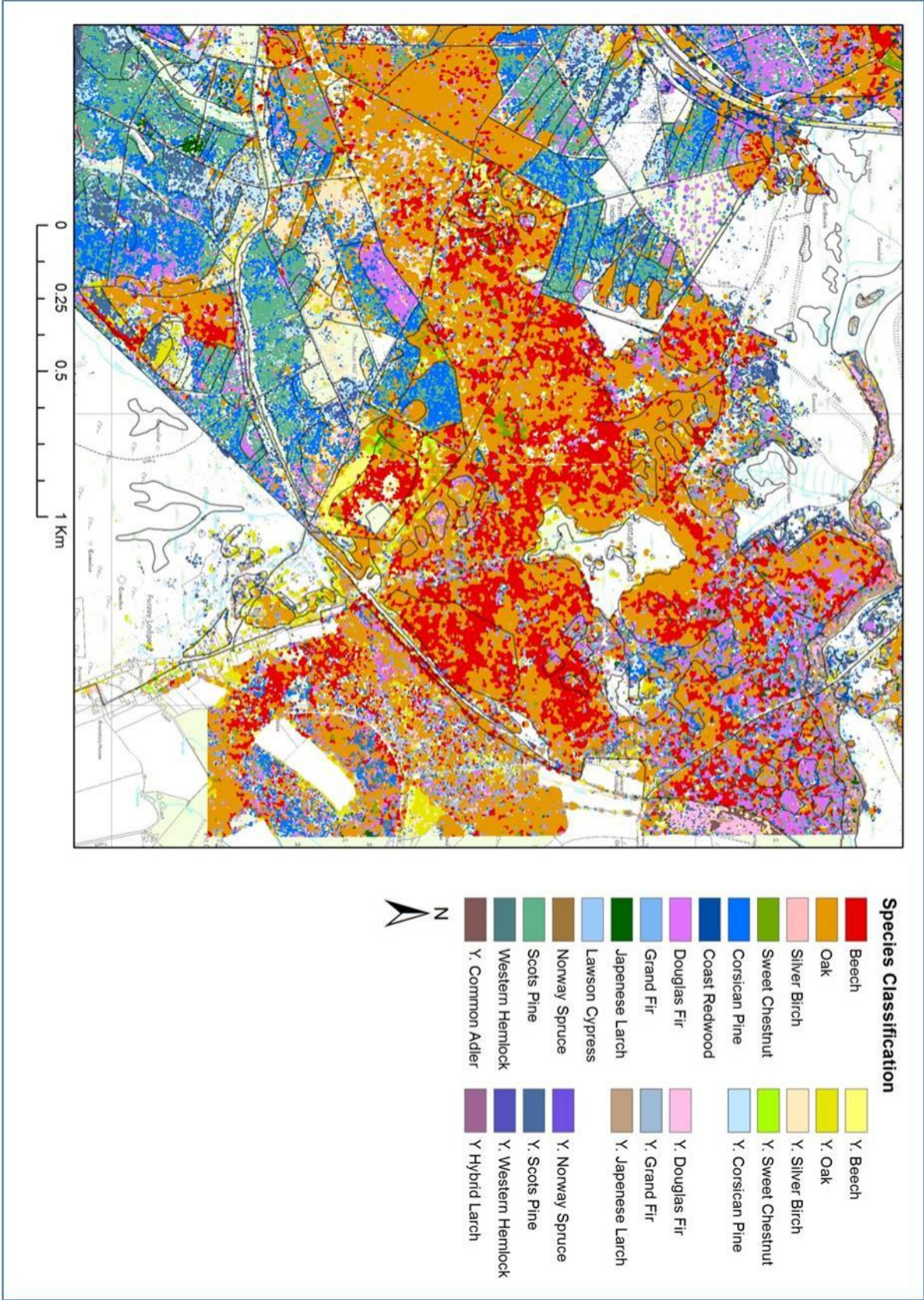


Figure 6.2 – Classification map part 2: Tantany and Frame Woods, and much of Hawkhill and part of Frame-Heath inclosures. (Note that “Y.<class_name>” denotes ‘young’, i.e. lower tree height and/or different hyperspectral signature than mature trees)

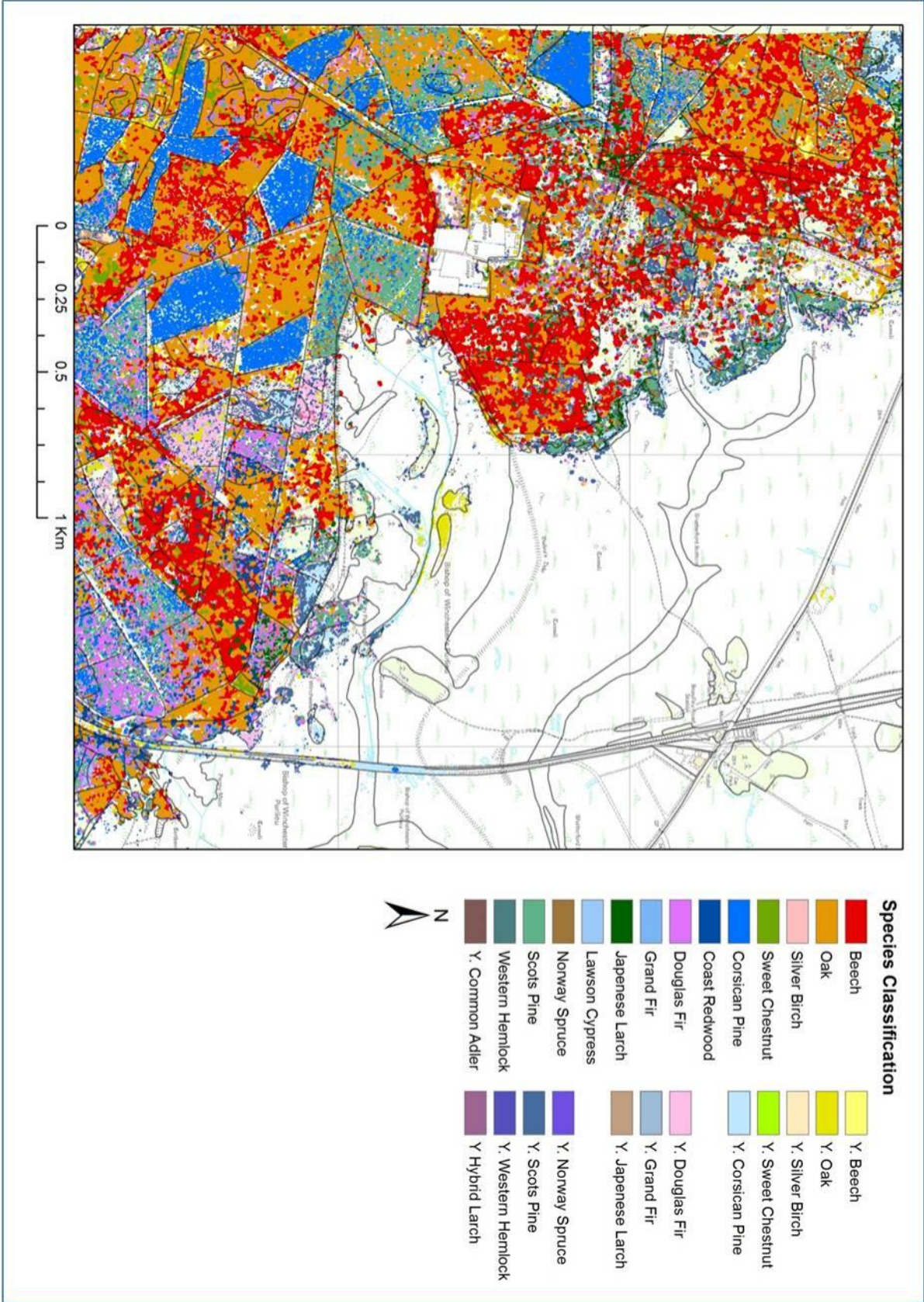


Figure 6.3 – Classification map part 3: Denny Lodge, Stubby Copse and Denny inclosures, in addition to Denny Wood. (Note that “Y.<class_name>” denotes ‘young’, i.e. lower tree height and/or different hyperspectral signature than mature trees)

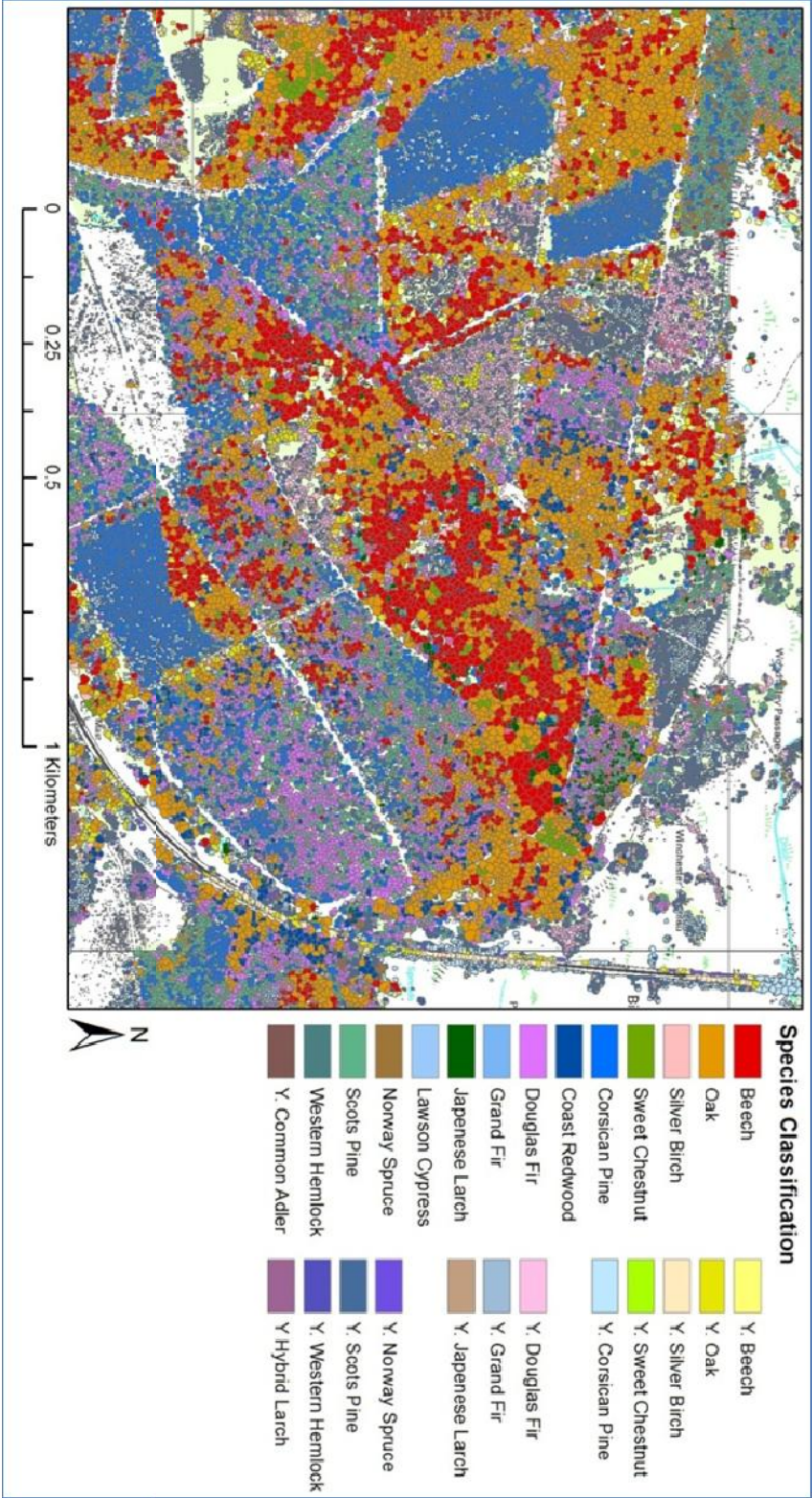


Figure 6.4 – A sample subset of the eCognition-derived classification map (Note that “Y.<class_name>” denotes ‘young’, i.e. lower tree height and/or different hyperspectral signature than mature trees)

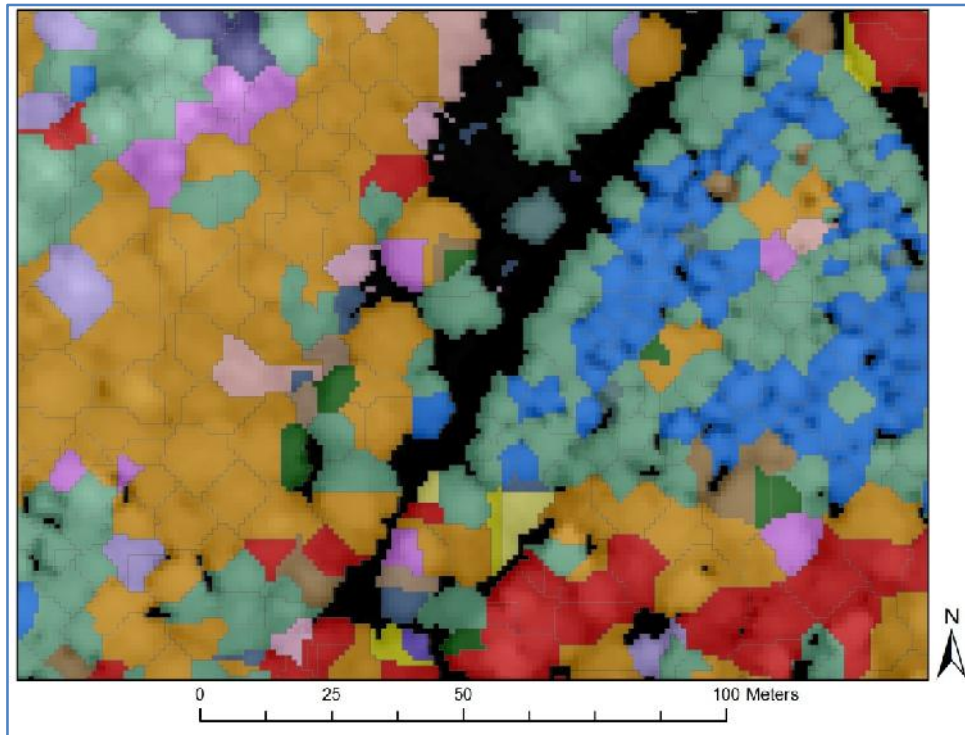


Figure 6.5 – Examples of over segmentation resulting in poor tree crown delineation illustrating both deciduous (left) and coniferous (right). The latter is generally less effected.

6.1.1 Assessment against Forestry Commission stand inventory information

To assess the remote sensing derived product against Forestry Commission (FC) stand inventory data, ten FC compartments were randomly selected from the supplied shapefile and compared against the species information from the intersecting classified tree crown objects. For the FC polygons only general species information is given in the form of up to three species and an estimate of the per-cent area coverage. Table 6.1 summarises the results from the ten FC polygons and species information from coincident derived tree crown objects. It should be noted that the values of species proportions for the remote sensing classification are for the cumulative area of each species class, a total calculated from the area in square metres of classified ITC objects, and converted into a percentage relative to the total area of ITC objects intersecting the FC polygon.

For each of the sample FC polygons, the ITC classification detects the dominant species types within the shapefile extents. The percentage cover of the remote sensing derived species information can differ by up to 40% from that in the FC data. A greater number of tree species types are detected within the polygons by the remote sensing classification, which may account for the disparity. The polygon boundaries for one compartment may also straddle other compartments.

It should be noted that the likelihood of there being only one-to-three tree species in the extent of one compartment is unlikely, a supposition generally supported by fieldwork results. This may cause large bias in the results presented in Table 6.1. For example, compartment number 4634 is reported by the FC as being 100% oak trees, whereas this is only accounted for by 62% of the ITC area classified as oak. Thus, one must question the validity of this comparison.

Commission and omission errors were calculated in addition to an estimate of map accuracy (MA), defined as follows (Short, 2005):

$$MA = \frac{X_{correct}}{X_{correct} + X_{omission} + X_{commission}} \quad [6.1]$$

For the ten sample FC compartment areas commission errors can vary from 25-51%, with an average of 32%. Omission error ranges from 15-48% with an average of 26%. Overall MA ranges from 52-67% with an average of 57%.

6.1.2 Assessment of classification against 2012 field validation data

This section presents the comparison of the field validation plots (recorded in 2012) against the remote sensing derived species classification ITC map, see Table 6.2 where the tree species and proportions in which they were encountered per plot are summarised. At the most basic level the number of species identified within an area corresponds closely between field data and the classified ITC map. The average remote sensing deviation is ± 1 with a maximum of 3 and a minimum of 0, against both total field-recorded stems and those classified as overstorey. The highest differences are encountered in plots of uniform species, i.e. plots 31 and 39.

Table 6.1 – Comparison between 10 FC compartments and coincident remote sensing derived classified objects.

FC compartment	FC Species and relative abundance:	Remote sensing derived species and relative abundance based on ITC objects:	Commission/Omission errors and Map accuracy
ID: 4632 (b) 436142 E; 102716 N	Scots pine: 55% Corsican Pine: 30% Not planted: 15%	Percentage cover: Corsican Pine 39.6% Scots Pine 20.1% Douglas Fir 6.7% Silver Birch 15.1% Oak 17.6%	<u>Commission:</u> 35% <u>Omission:</u> 27% <u>Map accuracy:</u> 52%
ID: 4634 (a) 435670 E; 102323 N	Oak: 100%	Percentage cover: Beech 21.8% Oak 60.5% Corsican Pine 3.8% Scots Pine 7.0% Douglas Fir 3.1% Silver Birch 3.6%	<u>Commission:</u> 26% <u>Omission:</u> 48% <u>Map accuracy:</u> 55%
ID: 4641 (a) 434679 E; 101985 N	Scots pine: 83% None: 17%	Percentage cover: Corsican Pine 22.1% Japanese Larch 2.2% Douglas Fir 2.9% Scots Pine 66.3%	<u>Commission:</u> 34% <u>Omission:</u> 22% <u>Map accuracy:</u> 54%
ID: 4624 (a) 434542 E; 101971 N	Oak: 100%	Percentage cover: Corsican Pine 2.1% Oak 62.1% Douglas Fir 5.0% Scots Pine 3.0% Grand Fir 3.8% Silver Birch 11.7% Japanese Larch 3.7% Sweet Chestnut 7.1%	<u>Commission:</u> 51% <u>Omission:</u> 15% <u>Map accuracy:</u> 58%
ID: 4626 (a) 434611 E; 102566 N	Corsican Pine: 100%	Percentage cover: Corsican Pine 76.1% Oak 2.9% Hybrid Larch 2.0% Scots Pine 17.1%	<u>Commission:</u> 25% <u>Omission:</u> 20% <u>Map accuracy:</u> 64%
ID: 4410 (a) 434570 E; 104637 N	Beech: 89% Oak: 11%	Percentage cover: Beech 59.7% Oak 27.4% Japanese Larch 4.3% Sweet Chestnut 4.8%	<u>Commission:</u> 34% <u>Omission:</u> 19% <u>Map accuracy:</u> 52%
ID: 4419 (a) 433660 E; 104220 N	Scots Pine: 46% Douglas fir: 35%	Percentage cover: Beech 3.5% Oak 5.3% Corsican Pine 30.3% Scots Pine 41.0% Douglas Fir 18.5%	<u>Commission:</u> 24% <u>Omission:</u> 43% <u>Map accuracy:</u> 64%
ID: 4439 (a) 433182 E; 104521 N	Beech: 71% Oak: 19% Birch: 9%	Percentage cover: Beech 47.4% Sweet Chestnut 2.8 % Oak 45.2	<u>Commission:</u> 33% <u>Omission:</u> 31% <u>Map accuracy:</u> 56%
ID: 4404 (a) 433349 E; 105139 N	Scots Pine: 96% Mixed broadleaves: 4%	Percentage cover: Beech 2.0% Oak 15.8% Corsican Pine 11.1% Scots Pine 59.8% Douglas Fir 6.2% Silver Birch 1.5% Norway Spruce 3.2%	<u>Commission:</u> 41% <u>Omission:</u> 19% <u>Map accuracy:</u> 56%
ID: 4406 (a) 433363 E; 104784 N	Oak: 50% Beech: 39%	Percentage cover: Beech 20.9% Silver Birch 3.9% Oak 70.5% Sweet Chestnut 2.3%	<u>Commission:</u> 25% <u>Omission:</u> 16% <u>Map accuracy:</u> 67%

The dominant tree species type can be detected using remote sensing methods in all but one of the validation field plots. Plot number 38, misclassified the coniferous species within its extent. There are examples of misclassifying species in all but two plots; typically this is for coniferous species. There are also a number of examples where deciduous species such as oak are misclassified.

ITC objects produced were fewer in number than tree stems in the field, typically less than half. Given the presence of suppressed trees encountered in a number of plots it is helpful to consider overstorey trees only; they are defined here as trees with a height equal to or greater than 15m.

The most noticeable examples of underestimation were in deciduous plots, for example plot 22 had a total of 20 stems (17 overstorey), whereas there were only 10 intersecting ITC objects. Similarly, plot 29 had a total of 30 stems (21 overstorey) and a total of 12 ITC objects. Extremely densely populated plots such as plots 23 and 30, suffered the greatest underestimation. For example, plot 23 received few ITC objects, of 71 stems (54 overstorey), there were only 32 ITC objects. Uniform composition plots, such as plots 31 and 39 fair slightly better. For example, plot 31 is composed of 24 stems and 22 ITC objects were located, whereas plot 39 has 16 stems and is overestimated with 23 ITC objects.

Commission and omission errors were calculated in addition to estimates of MA for the (i) correspondence between field data stem species counts and classified ITC objects, and (ii) the correspondence between field data relative total crown horizontal area of species types against the relative total crown area for classified ITC objects. For the former (i), the average MA was calculated as 48.7% from all 20 validation field plots for the correspondence of classified ITC objects against individual overstorey tree stems. For mainly deciduous plots MA was 51.2%, and for mainly coniferous plots MA was 48.3%. Average commission error for all 20 plots was 14%, and average omission error was 49%. For the latter (ii), the average MA was calculated as 62% from all validation field plots. For mainly deciduous plots the MA was 64%, and for mainly coniferous MA was 57%. Average commission error for all 20 plots was 27% and average omission error was 26%.

Table 6.2 –Comparison between fieldwork validation plots from 2012 and the remote sensing derived classification ITC map.(Overstorey defined as trees with a height >50% of the maximum within the plot)

Site no.	Dom. site species	Field plot data – tree species and number *(...) denote overstorey	Remote sensing derived species and number based on ITCs	Commission/ Omission errors and Map accuracy (ITC counts)	Commission/ Omission errors and Map accuracy (ITC areas)
22	Decid .	[Spec. No.: 5 (3)] Beech: 3 (2) Common Hawthorn: 2 (0) Midland Hawthorn: 1 (0) Oak: 8 (7) Silver Birch: 6 (6)	[Spec. No.: 2] Oak: 9 Scots Pine: 1	<u>Commission:</u> 22.2% <u>Omission:</u> 44.4% <u>Map accuracy:</u> 53.9%	<u>Commission:</u> 23% <u>Omission:</u> 23% <u>Map accuracy:</u> 62%
23	Conif .	[Spec. No.: 2 (1)] Hybrid Larch: 46 (44) Silver Birch: 25 (10)	[Spec. No.: 4] Corsican Pine: 4 Hybrid Larch: 22 Oak: 4 Silver Birch: 2	<u>Commission:</u> 17.5% <u>Omission:</u> 52.4% <u>Map accuracy:</u> 42.9%	<u>Commission:</u> 33% <u>Omission:</u> 33% <u>Map accuracy:</u> 49%
24	Decid .	[Spec. No.: 3 (3)] Beech: 47 (24) Corsican pine: 2 (2) Oak: 5 (4)	[Spec. No.: 4] Beech: 10 Corsican pine: 1 Oak: 2 Scots Pine: 1	<u>Commission:</u> 3.3% <u>Omission:</u> 56.7% <u>Map accuracy:</u> 41.9%	<u>Commission:</u> 13% <u>Omission:</u> 13% <u>Map accuracy:</u> 77%
25	Decid .	[Spec. No.: 3 (2)] Beech: 12 (10) Holly: 4 (0) Oak: 5 (5)	[Spec. No.: 3] Beech: 6 Norway Spruce: 1 Oak: 3	<u>Commission:</u> 6.7% <u>Omission:</u> 53.3% <u>Map accuracy:</u> 47.1%	<u>Commission:</u> 23% <u>Omission:</u> 23% <u>Map accuracy:</u> 63%
26	Decid .	[Spec. No.: 5 (3)] Beech: 4 (2) Douglas fir: 1 (1) Holly: 1 (0) Oak: 10 (10) Silver birch: 1 (0)	[Spec. No.: 4] Beech: 2 Douglas Fir: 1 Norway Spruce: 2 Oak: 4	<u>Commission:</u> 15.4% <u>Omission:</u> 46.15% <u>Map accuracy:</u> 46.7%	<u>Commission:</u> 21% <u>Omission:</u> 21% <u>Map accuracy:</u> 65%
27	Decid .	[Spec. No.: 3 (2)] Beech: 19 (8) Holly: 9 (0) Oak: 5 (5)	[Spec. No.: 3] Beech: 8 Oak: 1 Scots Pine: 1	<u>Commission:</u> 7.7% <u>Omission:</u> 30.1% <u>Map accuracy:</u> 64.3%	<u>Commission:</u> 1% <u>Omission:</u> 1% <u>Map accuracy:</u> 82%
28	Conif .	[Spec. No.: 3 (3)] Douglas Fir: 23 (22) Oak: 1 (1) Scots Pine: 11 (11)	[Spec. No.: 4] Corsican Pine: 1 Douglas Fir: 10 Oak: 2 Scots Pine: 7	<u>Commission:</u> 5.9% <u>Omission:</u> 47.1% <u>Map accuracy:</u> 51.4%	<u>Commission:</u> 9% <u>Omission:</u> 9% <u>Map accuracy:</u> 84%
29	Decid .	[Spec. No.: 2 (2)] Beech: 23 (14) Oak: 7 (7)	[Spec. No.: 2] Beech: 4 Oak: 8	<u>Commission:</u> 4.8% <u>Omission:</u> 47.6% <u>Map accuracy:</u> 52.2%	<u>Commission:</u> 19% <u>Omission:</u> 19% <u>Map accuracy:</u> 69%
30	Conif .	[Spec. No.: 5 (3)] Douglas Fir: 47 (40) Horse Chestnut: 1 (1) Scots Pine: 3 (0) Silver Birch: 36 (0) Sweet Chestnut: 3 (3)	[Spec. No.: 7] Corsican Pine: 3 Douglas Fir: 8 Japanese Larch: 2 Norway Spruce: 1 Oak: 2 Silver Birch: 5 Sweet Chestnut: 1	<u>Commission:</u> 17.0% <u>Omission:</u> 61.7% <u>Map accuracy:</u> 27.5%	<u>Commission:</u> 37% <u>Omission:</u> 30% <u>Map accuracy:</u> 48%
31	Conif .	[Spec. No.: 1 (1)] Douglas fir: 24 (24)	[Spec. No.: 3] Corsican Pine: 2 Douglas Fir: 13 Scots Pine: 7	<u>Commission:</u> 37.5% <u>Omission:</u> 45.8% <u>Map accuracy:</u> 39.4%	<u>Commission:</u> 20% <u>Omission:</u> 20% <u>Map accuracy:</u> 66%

Table 6.2 (continued)

Site no.	Dom. site species	Field plot data – tree species and number *(...) denote overstorey	Remote sensing derived species and number based on ITCs	Commission/ Omission errors and Map accuracy (ITC counts)	Commission/ Omission errors and Map accuracy (ITC areas)
32	Mixed	[Spec. No.: 4 (4)] Beech: 3 (1) Douglas fir: 7 (7) Oak: 5 (4) Scots Pine: 5 (5)	[Spec. No.: 4] Douglas Fir: 12 Corsican Pine: 2 Oak: 2 Scots Pine: 3	<u>Commission:</u> 41.2% <u>Omission:</u> 29.4% <u>Map accuracy:</u> 50.0%	<u>Commission:</u> 45% <u>Omission:</u> 45% <u>Map accuracy:</u> 38%
33	Conif.	[Spec. No.: 4 (4)] Birch: 3 (2) Corsican Pine: 2 (2) Lawson Cypress: 1 (1) Western Hemlock: 22 (22)	[Spec. No.: 4] Corsican Pine: 1 Oak: 1 Scots Pine: 10 Western Hemlock: 24	<u>Commission:</u> 51.9% <u>Omission:</u> 3.7% <u>Map accuracy:</u> 60.5%	<u>Commission:</u> 30% <u>Omission:</u> 28% <u>Map accuracy:</u> 54%
34	Mixed	[Spec. No.: 5 (4)] Corsican Pine: 10 (9) Douglas Fir: 6 (3) Norway Spruce: 2 (0) Oak: 8 (4) Silver Birch: 2 (1)	[Spec. No.: 4] Beech: 1 Corsican Pine: 10 Douglas Fir: 3 Scots Pine: 6	<u>Commission:</u> 47.0% <u>Omission:</u> 5.6% <u>Map accuracy:</u> 57.1%	<u>Commission:</u> 33% <u>Omission:</u> 35% <u>Map accuracy:</u> 49%
35	Conif.	[Spec. No.: 4 (4)] Beech: 1 (1) Coast redwood: 11 (11) Douglas Fir: 18 (17) Silver Birch: 3 (2)	[Spec. No.: 5] Beech: 1 Coast Redwood: 10 Douglas Fir: 2 Oak: 6 Scots Pine: 3	<u>Commission:</u> 30.0% <u>Omission:</u> 56.7% <u>Map accuracy:</u> 33.3%	<u>Commission:</u> 32% <u>Omission:</u> 30% <u>Map accuracy:</u> 52%
36	Conif.	[Spec. No.: 4 (4)] Douglas Fir: 1 (1) Oak: 5 (2) Scots Pine: 18 (18) Silver Birch: 1 (1)	[Spec. No.: 4] Beech: 2 Douglas Fir: 3 Oak: 3 Scots Pine: 13	<u>Commission:</u> 22.7% <u>Omission:</u> 22.7% <u>Map accuracy:</u> 63.0%	<u>Commission:</u> 35% <u>Omission:</u> 24% <u>Map accuracy:</u> 51%
37	Decid.	[Spec. No.: 3 (2)] Beech: 7 (4) Hawthorn: 1 (0) Oak: 15 (12)	[Spec. No.: 2] Coast Redwood: 1 Oak: 10	<u>Commission:</u> 6.2% <u>Omission:</u> 37.5% <u>Map accuracy:</u> 58.8%	<u>Commission:</u> 31% <u>Omission:</u> 31% <u>Map accuracy:</u> 52%
38	Mixed	[Spec. No.: 3 (3)] Beech: 7 (7) Oak: 6 (6) Western Hemlock: 23 (15)	[Spec. No.: 5] Beech: 8 Douglas Fir: 3 Norway Spruce: 1 Oak: 3 Scots Pine: 1	<u>Commission:</u> 21.4% <u>Omission:</u> 64.3% <u>Map accuracy:</u> 29.4%	<u>Commission:</u> 36% <u>Omission:</u> 36% <u>Map accuracy:</u> 47%
39	Conif.	[Spec. No.: 1 (1)] Douglas fir: 16 (15)	[Spec. No.: 4] Beech: 5 Corsican Pine: 3 Oak: 1 Douglas fir: 14	<u>Commission:</u> 60.0% <u>Omission:</u> 6.7% <u>Map accuracy:</u> 58.3%	<u>Commission:</u> 32% <u>Omission:</u> 32% <u>Map accuracy:</u> 51%
40	Decid.	[Spec. No.: 3 (3)] Beech: 11 (8) Douglas Fir: 2 (1) Oak: 4 (4)	[Spec. No.: 3] Beech: 2 Oak: 9 Sweet Chestnut: 1	<u>Commission:</u> 46.2% <u>Omission:</u> 52.9% <u>Map accuracy:</u> 45.8%	<u>Commission:</u> 42% <u>Omission:</u> 42% <u>Map accuracy:</u> 40%
41	Mixed	[Spec. No.: 5 (4)] Beech: 11 (3) Corsican Pine: 2 (2) Norway spruce: 1 (0) Oak: 4 (3) Scots Pine: 6 (6)	[Spec. No.: 5] Beech: 2 Corsican Pine: 4 Douglas Fir: 3 Oak: 4 Scots Pine: 3	<u>Commission:</u> 35.7% <u>Omission:</u> 35.7% <u>Map accuracy:</u> 50%	<u>Commission:</u> 20% <u>Omission:</u> 15% <u>Map accuracy:</u> 60%

6.2 Statistical models for estimating forest structural and compositional metrics from hyperspectral remote sensing

This section outlines the statistical models applied to the hyperspectral remote sensing metrics to estimate field plot-level metrics. A total of 155 remote sensing metrics were generated for each field plot extent and comprised summary metrics for both classified ITC objects and calculated VI products. Two statistical models, using Ordinary Least-Squares (OLS) and Akaike information criterion (AICc) multiple regression approaches, were calculated for each of the 39 field metrics using the inputs from the hyperspectral remote sensing dataset. Each of the coefficients and variables presented in this section were statistically significant ($p < 0.05$), and the fit of regression models to the data (field data collected in 2010) is indicated by R^2 . This section summarises the regression models produced by forest structural, forest compositional, deadwood, and regeneration and understorey metrics, and the two additional metrics required for condition index construction.

6.2.1 Forest structure within the plot

A listing of the 14 forest structural metrics and their associated R^2 values are given in Table 6.3, while the inputs to the regression models are listed in Table 6.4. When considering the statistical methods, OLS and AICc, in the majority of cases (nine of fourteen) AICc produced better R^2 values. Seven of the total fourteen field metrics have similar R^2 values (± 0.1) for AICc and OLS, the others can vary up to ± 0.2 . The greatest disparities are for the standard deviation of tree NN spacing, total plot basal area, percentage big trees (DBH > 40cm), average crown horizontal area, and standard deviation of crown horizontal area. In reference to Table 6.4 for structural metrics, where there is a higher R^2 value, the model invariably has more predictor inputs. For example the standard deviation of tree NN spacing has one input for OLS and three for AICc, with R^2 values of 0.571 and 0.812, respectively. Typically, AICc regression models utilise more predictor input variables than their OLS counterparts, for example average crown horizontal area has one input for OLS, and three for AICc. There is only one example of an OLS model using more predictor inputs; this was the percentage of big trees (DBH > 40cm).

AICc weights are presented in Table 6.3 (from the iterative regression model process), for the selected model, with the highest weight, and then the second highest weighted model. It should be noted that this model selection approach required many thousands of iterations. All delta AICc values for the selected model were 0, and all secondary models were $\Delta i < 2$. The AICc weight represents the ratio of delta AIC (Δi) values for each model relative to the whole set of candidate models computed within the R software. The interpretation of Akaike weights is straightforward: they indicate the probability that the model is the best among the whole set of candidate models. One can compare the Akaike weights of the “best” model and competing models to determine to what extent it is better than another, this is termed the evidence ratio. For example the ratio for ‘number of trees’ is 1.091, thus the model is therefore 1.091 times more likely than the second model to be the best, given the set of candidate models and the data. This suggests that the rank of the model might change if the researcher was to take a series of independent samples of identical size (Burnham and Anderson 2002). In other words, there would be a high degree of uncertainty regarding the best model in all cases.

In Table 6.4, the first four attributes utilise various component in their model construction. The first attribute, (1) number of trees, uses hyperspectral metrics relating to leaf pigment absorption and a broadband measure of vegetation greenness for both OLS and AICc models, the latter also uses the number of ITC objects detected. Both OLS and AICc (2) ‘average tree NN spacing’ models utilise metrics relating to narrow band vegetation greenness and canopy water content. The models for (3) ‘standard deviation of tree NN spacing’ differ for OLS and AICc, the former uses one measure of narrowband vegetation greenness, whilst the latter uses metrics related to the number of ITC objects, light use efficiency and broadband greenness. The fourth field metric, (4) ‘mean DBH’ is estimated by one narrowband greenness index for OLS, whereas AICc utilises a narrowband greenness index, the number of ITC objects and an estimate of canopy water content. A number of models include the SI index as a metric related to tree species diversity, for example the total basal area and standard deviation of the height of the first live branch.

Table 6.3 – Forest structural metrics, R^2 and AICc weights (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	OLS R^2	AICc R^2	AICc weight	AICc weight of 2 nd best	AICc evidence ratio
<i>Number of trees per 30x30m</i>	0.749	<u>0.808</u>	0.036	0.033	1.091
<i>Average tree spacing (m)</i>	0.482	<u>0.609</u>	0.024	0.024	1.000
<i>STDev. of tree spacing (m)</i>	0.571	<u>0.812</u>	0.143	0.116	1.233
<i>Mean DBH (cm)</i>	0.549	<u>0.691</u>	0.057	0.045	1.267
<i>STDev. of DBH (cm)</i>	0.768	<u>0.796</u>	0.055	0.047	1.170
<i>Basal area within 30x30m</i>	<u>0.753</u>	0.484	0.239	0.116	2.060
<i>Percentage big trees (DBH 40>x<80cm)</i>	0.505	<u>0.554</u>	0.075	0.047	1.596
<i>Percentage very big trees (DBH >80cm)</i>	<u>0.771</u>	0.744	0.029	0.029	1.000
<i>Percentage big trees (DBH >40cm)</i>	<u>0.735</u>	0.581	0.126	0.076	1.658
<i>Average height of the first live branch (m)</i>	<u>0.692</u>	0.672	0.075	0.063	1.190
<i>STDev. of the height of the first live branch (m)</i>	0.464	<u>0.588</u>	0.077	0.062	1.242
<i>Average crown horizontal area (m²)</i>	0.564	<u>0.879</u>	0.144	0.061	2.361
<i>STDev. crown horizontal area (m²)</i>	0.520	<u>0.727</u>	0.194	0.077	2.519
<i>Total crown horizontal area (m²)</i>	<u>0.820</u>	0.716	0.148	0.092	1.609

Table 6.4 – Hyperspectral metric inputs for each forest structural metric regression model (Appendix B.2 should be consulted for a description of hyperspectral metric names).

Metric Name:	OLS model input metrics	AICc model input metrics
<i>Number of trees per 30x30m</i>	(1) cri1_Sum (leaf-pigments) [Leaf-off] (2) sri2_Min (greenness) [Leaf-on] (3) sri2_STD (greenness) [Leaf-on]	(1) cri1_Sum (leaf-pigments) [Leaf-off] (2) no_tree (Direct – no. of ITC objects) (3) sri2_Min (greenness) [Leaf-on] (4) sri2_STD (greenness) [Leaf-on]
<i>Average tree spacing (m)</i>	(1) wbi1_STD (canopy water content) [Leaf-off] (2) rendvi2_Min (greenness) [Leaf-on]	(1) rendvi2_Min (greenness) [Leaf-on] (2) vrei2_STD (greenness) [Leaf-on] (3) wbi1_STD (canopy water content) [Leaf-off]
<i>STDev. of tree spacing (m)</i>	(1) mresri1_Max (greenness) [Leaf-off]	(1) no_tree (Direct – no. of ITC objects) (2) pri2_STD (light-use efficiency) [Leaf-on] (3) sri2_Range (greenness) [Leaf-on]
<i>Mean DBH (cm)</i>	(1) vrei1_STD (greenness) [Leaf-off]	(1) mresri2_STD (greenness) [Leaf-on] (2) no_tree (Direct – no. of ITC objects) (3) wbi1_Range (greenness) [Leaf-off]
<i>STDev. of DBH (cm)</i>	(1) mresri1_STD (greenness) [Leaf-off] (2) mrendvi1_Min (greenness) [Leaf-off]	(1) mrendvi1_STD (greenness) [Leaf-off] (2) mresri1_Sum (greenness) [Leaf-off] (3) sri1_Sum (greenness) [Leaf-off]
<i>Basal area within 30x30m</i>	(1) SI (Direct – from ITC objects) (2) arvi1_STD (greenness) [Leaf-off] (3) pri2_STD (light-use efficiency) [Leaf-on]	(1) arvi2_Min (greenness) [Leaf-on] (2) mrendvi1_Max (greenness) [Leaf-off]
<i>Percentage big trees (DBH 40>x<80cm)</i>	(1) rgri2_Range (light-use efficiency) [Leaf-on] (2) nat_spec (Direct – from ITC objects)	(1) arvi2_Min (greenness) [Leaf-on] (2) mresri2_STD (greenness) [Leaf-on] (3) pri2_Range (light-use efficiency) [Leaf-on]
<i>Percentage very big trees (DBH >80cm)</i>	(1) mresri1_STD (greenness) [Leaf-off] (2) rgri1_Range (light-use efficiency) [Leaf-off]	(1) mrendvi1_STD (greenness) [Leaf-off] (2) no_tree (Direct – from ITC objects)
<i>Percentage big trees (DBH >40cm)</i>	(1) rgri2_Range (light-use efficiency) [Leaf-on] (2) SI (Direct – from ITC objects) (3) vrei1_STD (greenness) [Leaf-off] (4) ari2_STD (leaf-pigments) [Leaf-on] (5) no_nat_t (Direct – from ITC objects)	(1) mresri2_STD (greenness) [Leaf-on] (2) no_tree (Direct – from ITC objects) (3) sri2_Min (greenness) [Leaf-on]
<i>Average height of the first live branch (m)</i>	(1) vrei1_Mean (greenness) [Leaf-off] (2) ari1_Range (leaf-pigments) [Leaf-off]	(1) nat_spec (Direct – from ITC objects) (2) rgri1_Sum (light-use efficiency) [Leaf-off]
<i>STDev. of the height of the first live branch (m)</i>	(1) ari2_Min (leaf-pigments) [Leaf-on] (2) SI (Direct – from ITC objects)	(1) SI (Direct from ITC objects) (2) sri2_Max (greenness) [Leaf-on] (3) sri2_Sum (greenness) [Leaf-on]
<i>Average crown horizontal area (m²)</i>	(1) mresri1_STD (greenness) [Leaf-off]	(1) mresri2_Max (greenness) [Leaf-on] (2) no_tree (Direct – from ITC objects) (3) sri2_STD (greenness) [Leaf-on] (4) wbi1_STD (canopy water content) [Leaf-off]
<i>STDev. crown horizontal area (m²)</i>	(1) mresri1_STD (greenness) [Leaf-off]	(1) ari2_Min (light-use efficiency) [Leaf-on] (2) no_tree (Direct – from ITC objects)
<i>Total crown horizontal area (m²)</i>	(1) pri2_Max (light-use efficiency) [Leaf-on] (2) arvi2_Min (greenness) [Leaf-on] (3) mresri2_Range (greenness) [Leaf-on]	(1) ari2_Sum (leaf-pigments) [Leaf-on] (2) pri2_Range (light-use efficiency) [Leaf-on] (3) SI (Direct – from ITC objects)

6.2.2 Forest composition within the plot

A listing of the seven forest compositional metrics and their associated R^2 values are given in Table 6.5, while the inputs to the regression models are listed in Table 6.6. Regression models could not be completed for three compositional metrics using OLS and one for AICc. For the former, the number of species, SH (by stem count) and SI (by basal area) failed, whereas only SI (by basal area) failed for AICc. It should also be noted that the number of species could be extracted directly, see section 6.3. When considering the statistical methods of OLS and AICc, three of the remaining four regression models in which a comparison can be made, had lower R^2 values for AICc. The difference was small however, <0.05 . The regression models for the SH index (by basal area) for the AICc had a higher R^2 value. Again, the AICc evidence ratio shows low ratio values for the selected model and the second ‘best’ model.

Table 6.6 indicates that various area-based facets of indices related to narrowband and broadband greenness, leaf-pigment canopy water content and light-use efficiency were used, from both leaf-on and leaf-off data. The first metric, the number of native trees, used the same metrics for both OLS and AICc regression models, which produced a different R^2 value due to different coefficients. The two inputs were indices of greenness from leaf on-data. SI (by stem count) uses different inputs when considering OLS and AICc models, the former uses greenness and canopy water content, while the latter uses light-use efficiency and canopy water content related variables, and also had a lower R^2 value. For the final metric, SH (by basal area), OLS has one input relating to ITC object intersecting the field plot extent classified as native species. The AICc equivalent was composed of four index-derived metrics relating to greenness and light use efficiency.

Table 6.5 – Forest composition metrics, R² and AICc weights (all models $p < 0.05$). Underlined values indicate highest R² value.

Metric Name:	OLS R ²	AICc R ²	AICc weight	AICc weight of 2 nd best	AICc evidence ratio
<i>Number of native trees per 30x30m</i>	<u>0.641</u>	0.623	0.033	0.032	1.031
<i>Percentage of native trees</i>	<u>0.752</u>	0.720	0.022	0.02	1.100
<i>No. of tree species per 30x30m</i>	-	<u>0.560</u>	0.065	0.043	1.512
<i>Shannon index (by stem count)</i>	-	<u>0.647</u>	0.059	0.058	1.017
<i>Simpson index (by stem count)</i>	<u>0.408</u>	0.345	0.089	0.064	1.391
<i>Shannon index (by BA)</i>	0.188	<u>0.620</u>	0.159	0.071	2.239
<i>Simpson index (by BA)</i>	-	-	-	-	-

Table 6.6 – Hyperspectral metric inputs for each forest compositional metric regression model (Appendix B.2 should be consulted for a description of hyperspectral metric names).

Metric Name:	OLS model input metrics	AICc model input metrics
<i>Number of native trees per 30x30m</i>	(1) sri2_Min (greenness) [Leaf-on] (2) vrei2_STD (greenness) [Leaf-on]	(1) sri2_Min (greenness) [Leaf-on] (2) vrei2_STD (greenness) [Leaf-on]
<i>Percentage of native trees</i>	(1) mresri1_Sum (greenness) [Leaf-off] (2) ari2_Min (pigment) [Leaf-on] (3) arvi1_Min (greenness) [Leaf-off]	(1) arvi1_Min (greenness) [Leaf-off] (2) sri1_Sum (greenness) [Leaf-off] (3) wbi2_STD (canopy water) [Leaf-on]
<i>No. of tree species per 30x30m</i>	-	(1) rgri1_Min (light-use efficiency) [Leaf-off] (2) sri2_STD (greenness) [Leaf-on] (3) wbi1_STD (canopy water) [Leaf-off]
<i>Shannon index (by stem count)</i>	-	(1) rgri1_Min (light-use efficiency) [Leaf-off] (2) rgri2_Min (light-use efficiency) [Leaf-on] (3) sri2_Range (greenness) [Leaf-on] (4) wbi1_Range (canopy water content) [Leaf-off] (5) wbi2_STD (canopy water content) [Leaf-on]
<i>Simpson index (by stem count)</i>	(1) rendvi1_Min (greenness) [Leaf-off] (2) wbi2_Mean (canopy water content) [Leaf-on]	(1) sipi2_STD (light-use efficiency) [Leaf-on] (2) wbi2_Range (canopy water content) [Leaf-on]
<i>Shannon index (by BA)</i>	(2) no_nat_t (direct – ITC objects of native class)	(1) ndvi2_STD (greenness) [Leaf-on] (2) pri2_Max (light-use efficiency) [Leaf-on] (3) rendvi2_Range (greenness) [Leaf-on] (4) sipi1_STD (light use efficiency) [Leaf-off]
<i>Simpson index (by BA)</i>	-	-

6.2.3 Deadwood within the plot

A listing of four forest deadwood metrics and their associated R^2 values are reported in Table 6.7, while the inputs to the regression models are listed in Table 6.8. All but one of the AICc models produced higher R^2 values than OLS, although the values were close (± 0.07). Regression results for standing deadwood volume and class had high R^2 values approaching or above 0.9. Downed deadwood R^2 values were lower however. The AICc evidence ratio shows that the selected model is at least two times better than the second potential regression model.

In Table 6.8, the most commonly used input metrics relate to either greenness of vegetation or the efficiency of light-use within the canopy. The first regression model for the volume of standing dead wood uses four inputs for OLS and two for AICc, the first two of these are the same in both models. The additional two inputs for OLS only yield an additional 0.07 increase in R^2 value. The standing deadwood decay class models differ in their inputs, although both utilise a measure of species richness or diversity. For OLS this is combined with metrics relating to canopy light-use efficiency, whereas AICc uses greenness. It should be noted that VI values for deciduous areas were different to those in coniferous areas, and standing deadwood was small or non-existent in coniferous areas. Both of these issues may cause issues with the validity of the regression model. Both downed deadwood volume and decay class models produced through OLS contain one input, whereas those of AICc contain two.

Table 6.7 – Forest deadwood metrics, R^2 and AICc weights (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	OLS R^2	AICc R^2	AICc weight	AICc weight of 2 nd best	AICc evidence ratio
<i>Standing dead wood volume (m^3)</i>	<u>0.967</u>	0.889	0.079	0.045	1.756
<i>Standing dead wood decay class</i>	0.885	<u>0.903</u>	0.078	0.065	1.200
<i>Downed dead wood volume (m^3)</i>	0.392	<u>0.548</u>	0.021	0.018	1.167
<i>Downed dead wood decay class</i>	0.335	<u>0.418</u>	0.061	0.035	1.743

Table 6.8 – Hyperspectral metric inputs for each forest deadwood metric regression model (Appendix B.2 should be consulted for a description of hyperspectral metric names).

Metric Name:	OLS model input metrics	AICc model input metrics
<i>Standing dead wood volume (m³)</i>	(1) mresri1_STD (greenness) [Leaf-off] (2) rgri2_Max (light-use efficiency) [Leaf-on] (3) pri1_Max (light-use efficiency) [Leaf-off] (4) arvi1_Min (greenness) [Leaf-off]	(1) mresri1_STD (greenness) [Leaf-off] (2) rgri2_Max (light-use efficiency) [Leaf-on]
<i>Standing dead wood decay class</i>	(1) rgri1_Range (light-use efficiency) [Leaf-off] (2) SH (Direct – from ITC objects) (3) pri2_Max (light-use efficiency) [Leaf-on]	(1) spec_no (Direct – from ITC objects) (2) sri1_Sum (greenness) [Leaf-off] (3) sri2_Range (greenness) [Leaf-on]
<i>Downed dead wood volume (m³)</i>	(1) rendvi1_Max (greenness) [Leaf-off]	(1) ari2_Max (leaf-pigment) [Leaf-on] (2) mrendvi1_Max (greenness) [Leaf-off]
<i>Downed dead wood decay class</i>	(1) sri2_Mean (greenness) [Leaf-off]	(1) rendvi2_STD (greenness) [Leaf-on] (2) sri2_STD (greenness) [Leaf-on]

6.2.4 Understorey metrics within the plot

A listing of the 12 forest understorey metrics and their associated R^2 values are given in Table 6.9, while the inputs to the regression models are listed in Table 6.10. There are a number of metrics here with high standard error values, most notably those related to the number of seedlings, and number of native seedlings. The number of seedlings could not be predicted using the AICc method. The AICc method overall produced higher R^2 values than OLS, except for three metrics, which were the number of native seedlings, the number of seedling species and the number of vascular species. The final two metrics, relating to the number of bryophyte species and the percentage of bare soil were identical, including the input metrics and coefficients. When considering the AICc evidence ratio the first two metrics relating to the number of saplings and the native saplings within the plot showed the highest ratio values. The ratio values for the other understorey metrics fall within the range of 1 to 2.

In Table 6.10, except for the last two metrics relating to the number of bryophyte species and percentage of bare soil, there is a great deal of difference between the inputs used. Canopy vegetation greenness, leaf-pigments and light-use efficiency are all used, in addition to the occasional use of direct estimates of overstorey species. The first four metrics listed in Table 6.10, illustrate the differences in input metric usage, where OLS models typically employ measures of greenness, while AICc models use leaf-pigment input metrics. The three AICc models with low R^2 values relating to seedling and vascular species contain only one or two input predictor metrics. The same can be said when comparing the R^2 values for the seedlings

SH index for native species, where the OLS value is much lower than the AICc counterpart. The OLS model uses two light-use efficiency and greenness inputs, whereas the AICc model uses four inputs: greenness, two metrics related to light-use efficiency and an estimate of overstorey species number.

Table 6.9 – Forest understorey metrics, R^2 , standard error and AICc weights (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	OLS R^2	AICc R^2	AICc weight	AICc weight of 2 nd best	AICc evidence ratio
<i>No. saplings per 30x30m</i>	0.597	<u>0.727</u>	0.346	0.149	2.322
<i>No. native saplings per 30x30m</i>	0.646	<u>0.678</u>	0.584	0.113	5.168
<i>Percentage of native saplings per 30x30m</i>	0.692	<u>0.911</u>	0.364	0.205	1.776
<i>No. of sapling species per 30x30m</i>	0.663	<u>0.798</u>	0.189	0.096	1.969
<i>No. seedlings per 30x30m</i>	<u>0.809</u>	-	-	-	-
<i>No. native seedlings per 30x30m</i>	<u>0.642</u>	0.310	0.146	0.081	1.802
<i>Percentage of native seedlings per 30x30m</i>	0.816	<u>0.856</u>	0.071	0.066	1.076
<i>No. of seedling species per 30x30m</i>	<u>0.860</u>	0.423	0.101	0.063	1.603
<i>Seedlings Shannon index for native species</i>	0.415	<u>0.735</u>	0.281	0.263	1.068
<i>No. vascular species</i>	<u>0.855</u>	0.437	0.043	0.032	1.344
<i>No. bryophyte species</i>	0.439	0.439	0.025	0.024	1.042
<i>Percentage of bare soil within 30x30m</i>	0.518	0.518	0.038	0.034	1.118

Table 6.10 – Hyperspectral metric inputs for each forest understorey metric regression model (Appendix B.2 should be consulted for a description of hyperspectral metric names).

Metric Name:	OLS model input metrics	AICc model input metrics
<i>No. saplings per 30x30m</i>	(1) mresri2_Max (greenness) [Leaf-on]	(1) ari2_Min (leaf-pigments) [Leaf-on] (2) pri1_Range (light-use efficiency) [Leaf-off]
<i>No. native saplings per 30x30m</i>	(1) mresri2_Max (greenness) [Leaf-on] (2) sipi2_STD (light-use efficiency) [Leaf-on] (3) arvi2_Range (greenness) [Leaf-on]	(1) ari2_Min (leaf-pigments) [Leaf-on] (2) pri1_Range (light-use efficiency) [Leaf-off]
<i>Percentage of native saplings per 30x30m</i>	(1) mresri2_STD (greenness) [Leaf-on]	(1) ari1_Mean (leaf-pigments) [Leaf-off] (2) cri1_Range (leaf-pigments) [Leaf-off] (3) mresri2_STD (greenness) [Leaf-on]
<i>No. of sapling species per 30x30m</i>	(1) sri2_Range (greenness) [Leaf-on] (2) ndvi2_STD (greenness) [Leaf-on] (3) SI (Direct – from ITC objects) (4) sipi1_STD (light-use efficiency) [Leaf-off]	(1) ari1_Mean (leaf-pigments) [Leaf-off] (2) cri1_Range (leaf-pigments) [Leaf-off] (3) mresri2_STD (greenness) [Leaf-on]
<i>No. seedlings per 30x30m</i>	(1) ari2_Min (leaf-pigments) [Leaf-on] (2) pri2_Sum (light-use efficiency) [Leaf-on] (3) Per_nat_tree2 (Direct – from ITC objects) (4) spec_no (Direct – from ITC objects)	-
<i>No. native seedlings per 30x30m</i>	(1) pri2_Sum (light-use efficiency) [Leaf-on] (2) spec_no (Direct – from ITC objects)	(1) wbi2_Max (canopy water content) [Leaf-on]
<i>Percentage of native seedlings per 30x30m</i>	(1) ari2_Min (leaf-pigments) [Leaf-on] (2) pri2_Max (leaf-pigments) [Leaf-on] (3) ari2_Mean (leaf-pigments) [Leaf-on]	(1) ari2_Min (leaf-pigments) [Leaf-on] (2) mresri2_Max (greenness) [Leaf-on] (3) wbi1_STD (canopy water content) [Leaf-off]
<i>No. of seedling species per 30x30m</i>	(1) mresri2_Max (greenness) [Leaf-on] (2) rgri1_Min (light-use efficiency) [Leaf-off] (3) spec_no (Direct – from ITC objects) (4) vrei2_STD (greenness) [Leaf-on]	(1) ari2_Min (leaf-pigments) [Leaf-on]
<i>Seedlings Shannon index for native species</i>	(1) pri2_Range (light-use efficiency) [Leaf-on] (2) arvi1_Min (greenness) [Leaf-off]	(1) arvi2_Range (greenness) [Leaf-on] (2) rgri1_Min (light-use efficiency) [Leaf-off] (3) sipi2_Range (light-use efficiency) [Leaf-on] (4) spec_no (Direct – from ITC objects)
<i>No. vascular species</i>	(1) sri2_Range (greenness) [Leaf-on] (2) mrendvi1_STD (greenness) [Leaf-off] (3) nat_spec (Direct – from ITC objects) (4) sipi1_Min (light-use efficiency) [Leaf-off]	(1) mresri2_Range (greenness) [Leaf-on] (2) vrei1_STD (greenness) [Leaf-off]
<i>No. bryophyte species</i>	(1) ari1_Min (leaf-pigments) [Leaf-off] (2) rgri2_Range (light-use efficiency) [Leaf-off]	(1) ari1_Min (leaf-pigments) [Leaf-off] (2) rgri2_Range (light-use efficiency) [Leaf-on]
<i>Percentage of bare soil within 30x30m</i>	(1) mresri2_Sum (greenness) [Leaf-on] (2) sipi1_STD (light-use efficiency) [Leaf-off]	(1) mresri2_Sum (greenness) [Leaf-on] (2) sipi1_STD (light-use efficiency) [Leaf-off]

6.2.5 Metrics required for condition index construction

The two forest metrics necessary for index construction and their associated R^2 values are given in Table 6.11, while the inputs to the regression models are listed in Table 6.12. These two metrics could not be modelled by OLS regression, therefore only AICc model summaries are reported. The AICc models produced are very similar to the second best model. The input metrics in the model for the Sum of the DBH differences between neighbouring trees are indices related to greenness, the number of overstorey tree objects and a measure of leaf-pigments. For the second metric, the index of vertical separation, inputs are composed entirely of measures of greenness.

Table 6.11 – Condition index construction metrics, R^2 , standard error and AICc weights (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	OLS R^2	AICc R^2	AICc weight	AICc weight of 2 nd best	AICc evidence ratio
Sum of the DBH differences between neighbouring trees (cm)	-	<u>0.815</u>	0.043	0.043	1
Index of vertical separation	-	<u>0.547</u>	0.044	0.044	1

Table 6.12 – Hyperspectral metric inputs for each forest condition index metric regression model (Appendix B.2 should be consulted for a description of hyperspectral metric names).

Metric Name:	OLS model input metrics	AICc model input metrics
Sum of the DBH differences between neighbouring trees (cm)	-	(1) cri1_Sum (leaf-pigments) [Leaf-off] (2) no_tree (Direct – from ITC objects) (3) sri2_Min (greenness) [Leaf-on] (4) sri2_STD (greenness) [Leaf-on]
Index of vertical separation	-	(1) mresri2_STD (greenness) [Leaf-on] (2) ndvi2_Min (greenness) [Leaf-on]

6.3 Validation of hyperspectral remote sensing derived model estimates against field data

This section outlines the various direct and statistical model estimates in comparison with validation field data. It begins with the direct comparisons which can be made via remote sensing data, principally individual tree number. This is followed by a summary of the various statistical model estimates derived from remote sensing and their relation to field validation measurements. RMSE and NRMSE values are presented to illustrate the accuracy of each model prediction. NRMSE is a modified RMSE value to account for the field data population range, and produces a value of 0 to 1. NRMSE values of 0.5 or greater indicate poor model performance.

6.3.1 Direct comparisons between field measurements and hyperspectral remote sensing metrics

Table 6.13 summarises the relative proportions of native tree species detected within the field plot extents. Since actual population counts may not be accurate, due to over-/under-segmentation, a relative proportion was calculated as a percentage. Overall RMSE was calculated as 22.47% with a NRMSE of 0.22. Likewise the Shannon and Simpson indices of diversity were calculated from classified hyperspectral data, this is summarised in Table 6.14. For the Shannon index the RMSE was 0.56 and NRMSE was 0.39. For the Simpson index the RMSE was 0.28 and NRMSE was 0.22.

A number of additional metrics were required for the construction of the complex stand diversity index (CSDI). The CSDI is composed of four component indices relating to different forest stand attributes. The first of the CSDI component indices requires inputs from metrics calculated from the classified ITC hyperspectral map. This index relates to the total number of tree species and the relative proportions of the most and least abundant tree species. Table 6.15 summarises tree species number and maximum and minimum relative abundance. The disparity between remote sensing estimated and field recorded tree species numbers is relatively low. RMSE for estimates of the maximum population number are high, with a tendency to underestimate the field value. Likewise minimum population estimates also underestimate field recorded values, but given the smaller range of values the difference is less noticeable.

Table 6.13 – The proportion of native species according to classification

Site no.	Field native proportion (%)	eCognitionnative tree proportion (%)
22	100.00	100.00
23	35.21	21.21
24	98.11	92.86
25	100.00	90.00
26	94.12	60.00
27	100.00	90.91
28	34.29	42.86
29	100.00	100.00
30	46.67	31.82
31	0.00	31.00
32	65.00	26.32
33	10.71	30.56
34	46.43	35.00
35	12.12	43.48
36	96.00	81.82
37	100.00	90.91
38	36.11	75.00
39	100.00	60.87
40	88.24	91.67
41	91.67	61.11
RMSE		22.47
NRMSE		0.22

Table 6.14 – Comparison of diversity indices produced from remote sensing data and field data

Site no.	Field Shannon index (by stem no.)	Hyper. Shannon index	Field Simpson index (by stem no.)	Hyper. Simpson index
22	1.39	0.33	0.75	0.18
23	0.65	1.06	0.46	0.52
24	0.46	0.90	0.24	0.46
25	0.98	0.90	0.61	0.54
26	1.15	1.47	0.63	0.74
27	0.96	0.89	0.59	0.45
28	0.74	1.23	0.48	0.65
29	0.54	0.64	0.37	0.44
30	0.98	1.83	0.57	0.81
31	0.00	0.73	0.00	0.43
32	1.35	1.06	0.77	0.55
33	0.74	0.83	0.38	0.48
34	1.43	1.14	0.76	0.64
35	1.02	1.46	0.60	0.71
36	0.82	1.21	0.46	0.60
37	0.78	0.3	0.50	0.17
38	0.90	1.32	0.54	0.67
39	0.00	1.45	0.00	0.74
40	0.87	0.72	0.54	0.40
41	1.41	1.54	0.50	0.77
	Hyper. Shannon index		Hyper. Simpson index	
RMSE	0.56		0.28	
NRMSE	0.39		0.22	

Table 6.15 – A summary of input metrics for the first component index of the CSDI relating to relative species abundance. (Field = field data; RS = remote sensing data)

Site no.	Species no. [Field]	Species no. [RS]	Max. pop. [Field]	Max. pop. [RS]	Min. pop. [Field]	Min. pop. [RS]
22	4	4	8	8	1	1
23	2	4	46	22	25	1
24	3	4	47	10	2	1
25	3	4	12	6	4	1
26	5	3	10	4	1	1
27	3	5	19	8	5	1
28	3	4	23	10	1	1
29	2	4	23	8	7	4
30	5	4	47	7	1	1
31	1	4	24	12	0	2
32	4	3	7	12	3	2
33	4	1	22	24	1	1
34	5	4	10	10	2	1
35	4	3	18	10	1	1
36	4	4	18	13	1	1
37	4	5	15	10	1	1
38	3	4	23	8	6	1
39	1	2	16	8	0	1
40	3	4	11	9	2	1
41	5	5	11	6	1	2
-	RMSE	1.414	RMSE	15.392	RMSE	5.679
-	NRMSE	0.354	NRMSE	0.385	NRMSE	0.227

6.3.2 Validation of hyperspectral remote sensing derived statistical models against field data

This section contains the results of the regression models derived from hyperspectral remote sensing data compared to 2012 field data measurements. RMSE and NRMSE were calculated for each of the hyperspectral data derived models in order to identify which of the regression models best estimated reality. An overview of the OLS and AICc regression model predictions validated against field data is presented in Tables 6.16 and 6.17, for overstorey and understorey metrics respectively. In addition six metrics derived from the ITC classified image are presented in Table 6.16 for comparison, these were the number of trees, the number of native trees, the percentage of native trees, the number of tree species, the ITC derived Shannon index and the ITC derived Simpson index.

Two field-level metrics could not be adequately modelled by either regression method; these were SI index (by basal area) and the number of bryophyte (moss) species. A number of field metrics could only be modelled by one of the statistical approaches. Those modelled only by OLS are the number of sapling species and the number of seedlings per 30x30m, whereas that only modelled by AICc was the number of tree species.

In reference to both Tables 6.16 and 6.17, ten of the metrics produced poor RMSE and NRMSE results, arbitrarily considered here to be NRMSE values over 0.5. Four of these metrics relate to overstorey and understorey composition, whereas the remaining six relate to various structural metrics. The statistically derived estimates of the number of tree species are very poor; however the direct count of tree species provided from the classification is much better. The assessment of hyperspectral derived regression model estimates here also shows that twelve forest structural metrics, such as average height to the first live branch and the total crown area, produced good (NRMSE <0.4) estimates of validation field data.

Table 6.16 – The OLS and AICc hyperspectral regression models. Part 1: overstorey composition.

Field metric	ITC Direct RMSE*	ITC Direct NRMSE*	OLS RMSE	OLS NRMSE	AICc RMSE	AICc NRMSE
<i>Number of trees per 30x30m</i>	<u>22.659</u>	<u>0.306</u>	31.468	0.425	29.775	0.402
<i>Number of native trees per 30x30m</i>	15.987	0.307	15.676	0.301	<u>15.636</u>	<u>0.301</u>
<i>Percentage of native trees</i>	<u>30.050</u>	<u>0.300</u>	33.022	0.330	36.474	0.365
<i>No. of tree species per 30x30m</i>	<u>1.830</u>	<u>0.305</u>	-	-	6.047	1.008
<i>Average tree spacing (m)</i>	-	-	6.573	1.979	6.344	1.910
<i>STDev. Of tree spacing (m)</i>	-	-	1.142	1.375	1.101	1.325
<i>Mean DBH (cm)</i>	-	-	23.098	0.793	<u>14.503</u>	<u>0.498</u>
<i>STDev. of DBH (cm)</i>	-	-	<u>16.697</u>	<u>0.342</u>	22.858	0.469
<i>Basal area within 30x30m</i>	-	-	1.550	0.483	<u>1.414</u>	<u>0.441</u>
<i>(i) Percentage big trees (DBH 40>x<80cm)</i>	-	-	<u>25.824</u>	<u>0.369</u>	31.199	0.446
<i>(ii) Percentage very big trees (DBH >80cm)</i>	-	-	13.374	2.207	10.330	1.704
<i>(iii) Percentage big trees (DBH >40cm)</i>	-	-	44.407	0.634	36.286	0.518
<i>Shannon index (by stem count)</i>	<u>0.560</u>	<u>0.380</u>	-	-	0.669	0.467
<i>Simpson index (by stem count)</i>	<u>0.280</u>	<u>0.220</u>	0.420	0.326	0.519	0.403
<i>Shannon index (by BA)</i>	-	-	0.453	0.589	0.737	0.959
<i>Simpson index (by BA)</i>	-	-	-	-	-	-
<i>Average height of the first live branch (m)</i>	-	-	4.115	0.380	<u>3.446</u>	<u>0.318</u>
<i>STDev. of the height of the first live branch (m)</i>	-	-	2.090	0.405	<u>2.030</u>	<u>0.394</u>
<i>Average crown horizontal area (m²)</i>	-	-	43.783	0.503	70.562	0.810
<i>STDev. crown horizontal area (m²)</i>	-	-	27.243	0.476	<u>25.377</u>	<u>0.444</u>
<i>Total crown horizontal area (m²)</i>	-	-	568.027	0.393	<u>538.763</u>	<u>0.373</u>

*Denotes that this was not derived statistically and cannot be applied to all metrics

Table 6.17 – The OLS and AICc hyperspectral regression models. Part 2: understorey composition.

Field metric	OLS RMSE	OLS NRMSE	AICc RMSE	AICc NRMSE
Standing dead wood volume (m ³)	5.422	0.656	<u>3.915</u>	<u>0.474</u>
Standing dead wood decay class	0.269	0.403	<u>0.238</u>	<u>0.357</u>
Downed dead wood volume (m ³)	6.911	0.759	6.467	0.711
Downed dead wood decay class	<u>0.159</u>	<u>0.339</u>	0.219	0.467
No. saplings per 30x30m	<u>24.418</u>	<u>0.226</u>	25.247	0.234
No. native saplings per 30x30m	7.442	0.196	<u>6.761</u>	<u>0.178</u>
Percentage of native saplings per 30x30m	78.581	0.786	77.872	0.779
No. of sapling species per 30x30m	1.118	0.373	-	-
No. seedlings per 30x30m	362.970	0.388	-	-
No. native seedlings per 30x30m	239.764	0.333	<u>226.335</u>	<u>0.314</u>
Percentage of native seedlings per 30x30m	<u>43.431</u>	<u>0.434</u>	47.492	0.475
No. of seedling species per 30x30m	4.658	0.358	<u>4.105</u>	<u>0.316</u>
Seedlings Shannon index for native species	0.890	0.552	<u>0.867</u>	<u>0.537</u>
No. vascular species	4.327	0.721	3.491	0.582
No. bryophyte species	-	-	-	-
Percentage of bare soil within 30x30m	47.157	0.496	47.157	0.496
Sum of the DBH differences between neighbouring trees (cm)	-	-	3.271	0.409
Index of vertical separation	-	-	0.303	0.600

6.4 Summary of findings

The object-based classification produced a total of 26 tree species classes, including young and mature trees. A number of issues were encountered however. The first of these related to the over-segmentation of tree crowns, i.e. creating more than one object for a single tree crown, which could be miss-classified. The total number of tree crown objects did not typically represent the total number of tree stems within a sample plot, although a better comparison was achieved with overstorey trees.

The FC inventory data only provided up to three species classes which were present within that stand along with their relative proportions as a percentage. Overall map accuracy of the object-based classification against the FC species classes was 57% using the relative total area of those classified objects which intersected with the FC compartment boundaries. It

should be noted however that the field data often identified that more than three species were present indicating that some error was due to the FC data itself.

An assessment of the object-based classification was performed against field data in two ways. The first was a comparison of the species proportions of overstorey stem species (i.e. those trees larger than 15m tall) and the second was in terms of the relative total horizontal crown area of the identified species types. For the former, average species classification mapping accuracy was calculated as 49% comparing counts of classified ITC objects to the validation field plot counts of individual trees. For the latter, the mapping accuracy increased to 62% when comparing the relative total horizontal crown area of species for field and ITC objects by species type. The latter is probably the most valid approach.

Regression models could be created for all but two of the total thirty-nine indicator metrics, SI index (by basal area) and number of bryophyte species. Four OLS regression models and two AICc models did not produce acceptable results. Of the remainder, thirteen models produced poor results, based on RMSE and NRMSE values, for both OLS and AICc models. The number of species, SI and SH indices were better predicted directly and not through statistical means. OLS methods performed better for seven of the remaining twenty-three models. NRMSE is a modified RMSE value to account for the field data population range, and produces a value of 0 to 1. NRMSE values of 0.5 or greater indicate poor model performance.

Those models predicted most accurately (NRMSE <0.35), were the number of native trees, the percentage of native trees, the standard deviation of DBH, DDW decay class, number of saplings, number of native saplings, number of seedlings and number of seedling species.

Direct estimates of SH and SI indices of diversity and the number of tree species performed better than those derived statistically. The CSDI had four component indices, one of which relates to tree species abundance. This component index derived from the classified ITC map underestimated both maximum and minimum populations.

Chapter 7 – Forest assessment using LiDAR data

LiDAR products are assessed within this section in relation to field measurements, in addition to a broad comparison of the outputs of DR and FW LiDAR systems. The section begins with a comparison of LiDAR DTM products with ground GPS and total station measurements outside and underneath the canopy. A basic comparison of DR and FW datasets is presented for three sample flight lines. The remaining sections document the statistical models for forest field variable estimation created through the two regression approaches (OLS and AICc), first for DR and then for FW datasets.

7.1 Assessment of DR and FW LiDAR pre-processing data products

7.1.1 Accuracy assessment of DTMs against GPS ground control

This section concerns the accuracy assessment of the LiDAR derived elevation models, or more specifically ground-level. A total of 95 dGPS and total station measurements were taken in 2010 coincident with the fieldwork campaign at ground-level, both outside the forest and under forest canopy. These measurements consisted of Easting (X), Northing (Y) and elevation (Z) using the British National Grid coordinate system and the OS GB 1936 Datum. For the same coordinates elevation values were extracted from DTMs produced from DR and FW LiDAR under both leaf-on and leaf-off conditions, resulting in four datasets. Table 7.1 reports the calculated RMSE and NRMSE of the DTMs in relation to the field data. All LiDAR derived DTMs have an RMSE of <2m, where FW estimates have slightly smaller RMSE, by approximately 0.06m. Both leaf-off RMSE values are marginally smaller than their leaf-on counterparts.

Table 7.1 – RMSE and NRMSE of LiDAR DTMs in relation to ground GPS and total station measurements for all 95 values.

	DR Leaf-on (m)	DR Leaf-off (m)	FW Leaf-on (m)	FW Leaf-off (m)
All points				
RMSE	1.879	1.784	1.823	1.763
NRMSE	0.075	0.071	0.073	0.070

7.1.2 Comparison of DR and FW small-footprint LiDAR outputs

This section summarises the observed differences in the DR and FW LiDAR datasets for a sample of forested areas. It should be reiterated that the pulse spacing for DR and FW datasets was not equal resulting in FW having a lower pulse density for both data acquisition dates. For the purposes of comparison three sample flight lines were selected from the leaf-on 2010 datasets, corresponding to flight lines 11, 12 and 16, and subset to include only forested areas, as illustrated in Figure 7.1. Non-overlapping flight lines were chosen to avoid the complication of double sampling in the readout statistics, in addition to their coverage of different forest structural types.

The DR data had a maximum of four returns per pulse whereas the FW data could have up to ten returns per pulse which were generated through applying Gaussian curve fitting to each of the returned laser waveforms. Table 7.2 summarises the total number of emitted pulses and derived points for both DR and FW datasets by flight line subset after pre-processing. The total number of initial FW pulses were equivalent to approximately 40% of total number of initial DR pulses. From the values in Table 7.2, the difference in initial pulses compared with the total number of points for DR LiDAR only represents, on average, a 30% increase, whereas for FW LiDAR there is, on average, a 143% increase in the number of points relative to the initial number of pulses.

It should be noted that the different pulse rate between the DR and FW LiDAR datasets influenced the 3D spatial sampling and point density within the forest canopy. DR LiDAR had a higher density of spatial sampling overall, whereas the FW LiDAR had more points vertically; however there was a larger horizontal spacing at all vertical levels.

To compare the ability of the LiDAR systems to record points beneath the forest canopy, four sample areas were created in each of the three flight line extents. These twelve sample areas were 100m by 20m in size and the corners were oriented north to south. Figure 7.2 illustrates their locations. A number of statistics were calculated for both DR and FW LiDAR data subset areas. The number of points located at ten vertical levels (height-bins) was calculated. The ten height-bins were delineated as (i) 0-3m, (ii) 3-6m, (iii) 6-9m, (iv) 9-12m, (v) 12-15m, (vi) 15-18m, (vii) 18-21m, (viii) 21-24m, (ix) 24-27m and (x) 27m and higher. Figures 7.3-7.5 summarise the point totals for each of these height bins located within the sample areas. Figure 7.3 relates to flight line 11, and indicates that FW provides a higher number of points

within the mid canopy for both deciduous and coniferous sites, at approximately 9-18m. High numbers of points for both DR and FW appear in the 0-3m height bin, but DR usually has a higher return count. This pattern is repeated in Figure 7.4, which relates to flight line 12. Sample site 3 (line 12), a conifer stand, shows FW returns drastically outnumber the DR counterparts. In Figure 7.5, which relates to flight line 16, the same pattern is repeated, although DR counts can exceed FW in the upper canopy. For sample site 1 (line 16), a deciduous stand, FW produces larger height return counts in the mid canopy, (9-15m). Overall however there are no systematic differences within the overall distributions of LiDAR points between DR and FW datasets.

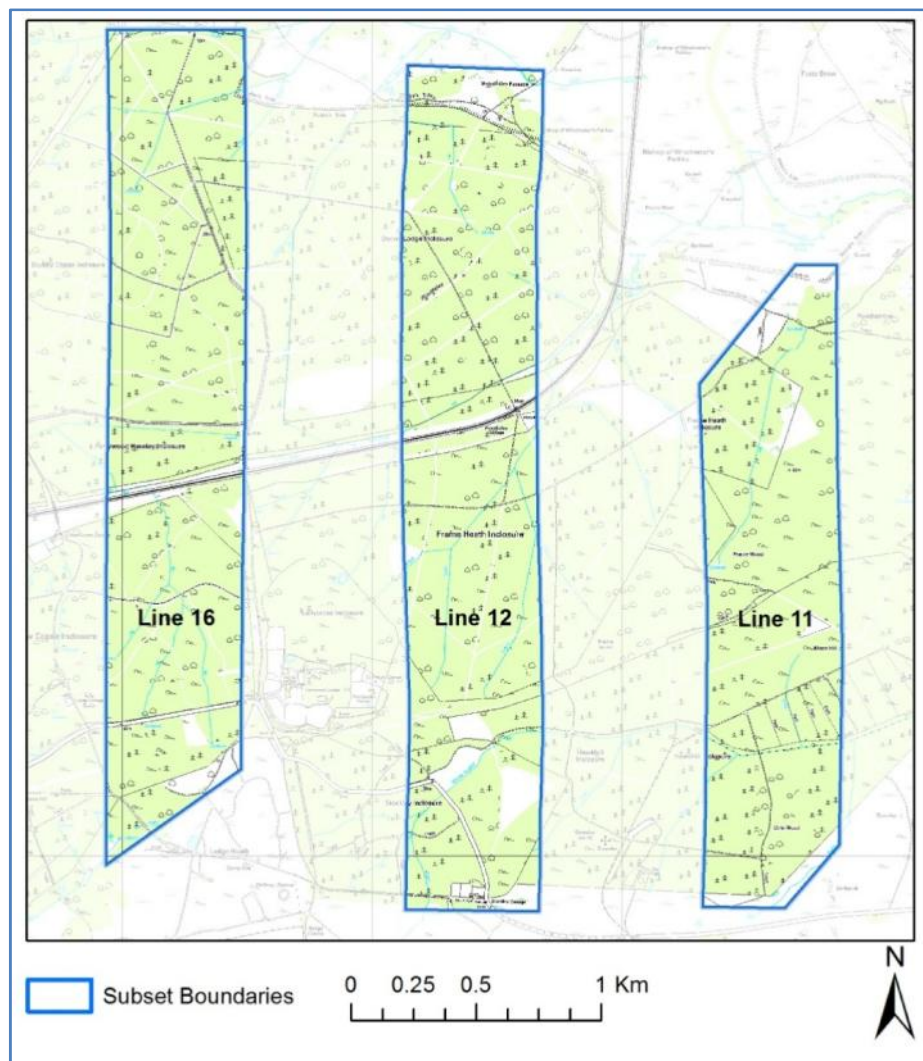


Figure 7.1 – An illustration of the locations of flight lines 11, 12 and 16 for 2010 leaf-on LiDAR data acquisitions. Base Map layer is © Crown Copyright/database right 2010. An Ordnance Survey/EDINA supplied service.

Table 7.2 – Summary of DR and FW number of recorded pulses and points by flight line subset.

Flight line no.	Extent (km ²)	DR total no. pulses	DR total no. points	FW total no. pulses	FW total no. points
11	1.292	4706392	6146380	3083695	7267835
12	1.838	12063215	15816682	4790090	11469291
16	1.737	11376678	14912713	4360583	11028303

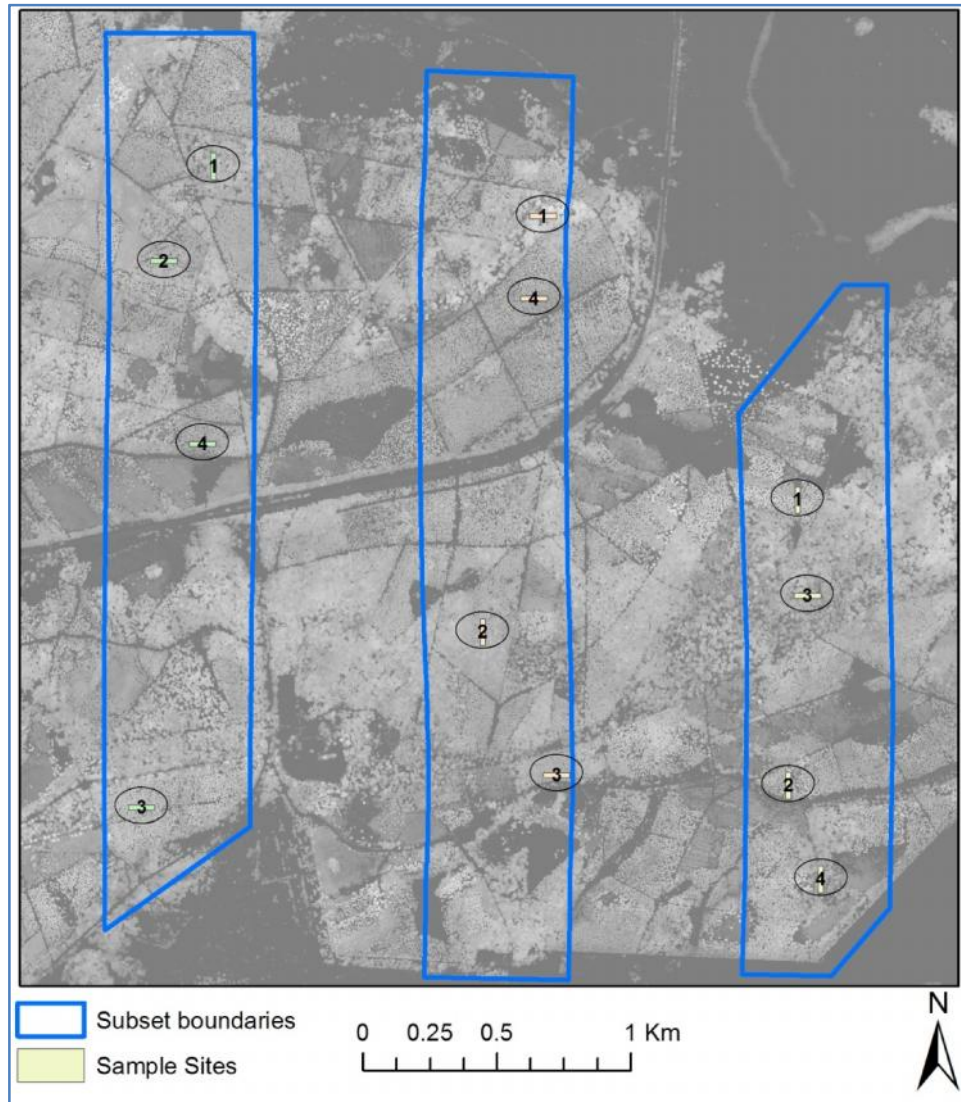


Figure 7.2 – The locations of samples sites

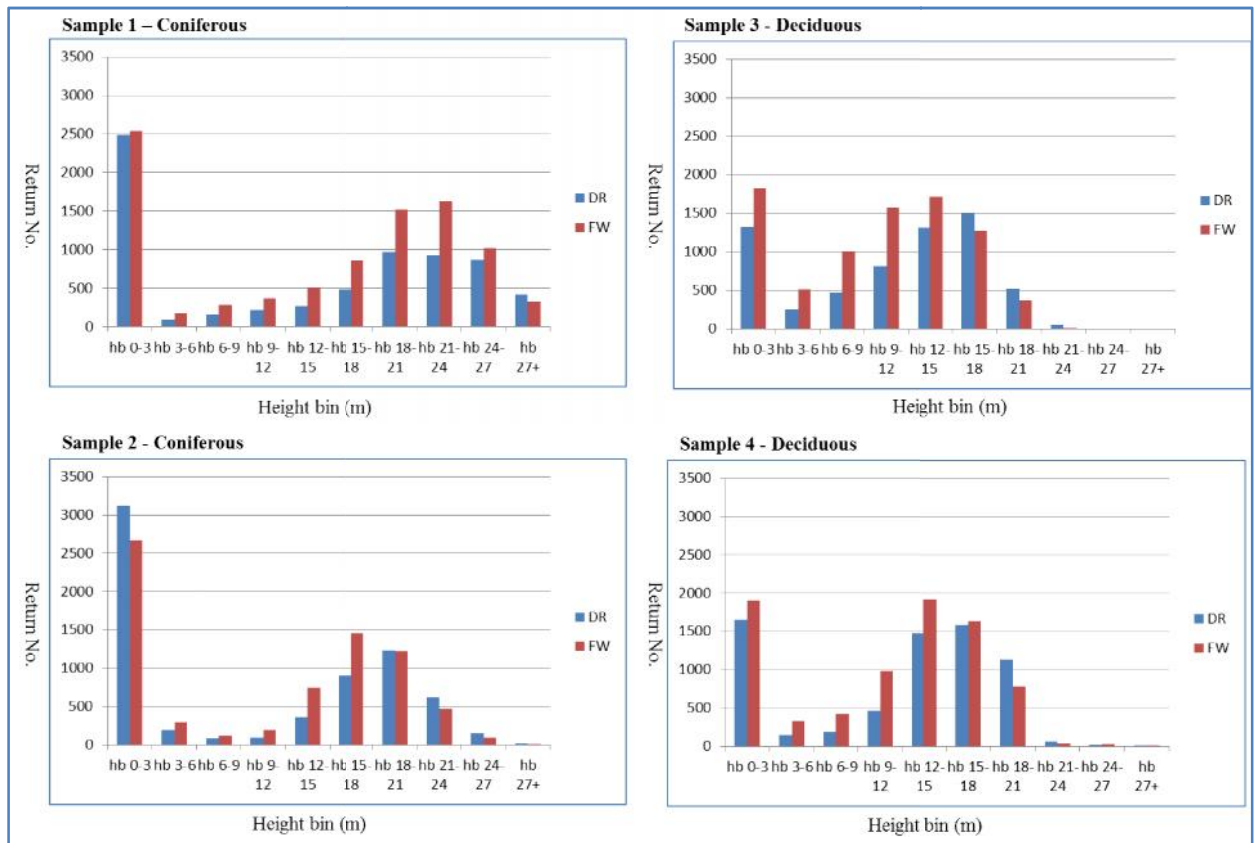


Figure 7.3 – Summary of height bin point counts for each of the sample areas in flight line 11 (both DR - blue, and FW - red)

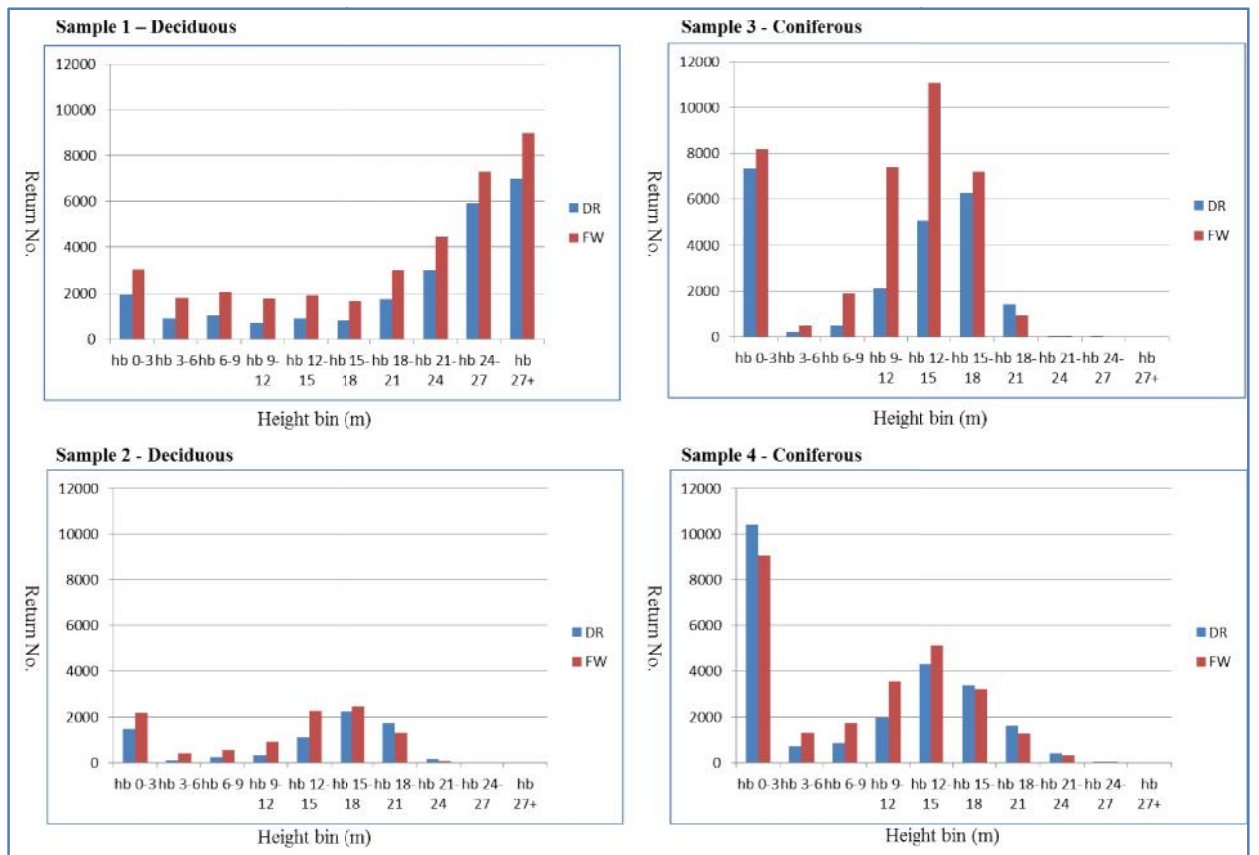


Figure 7.4 – Summary of height bin point counts for each of the sample areas in flight line 12 (both DR - blue and FW - red)

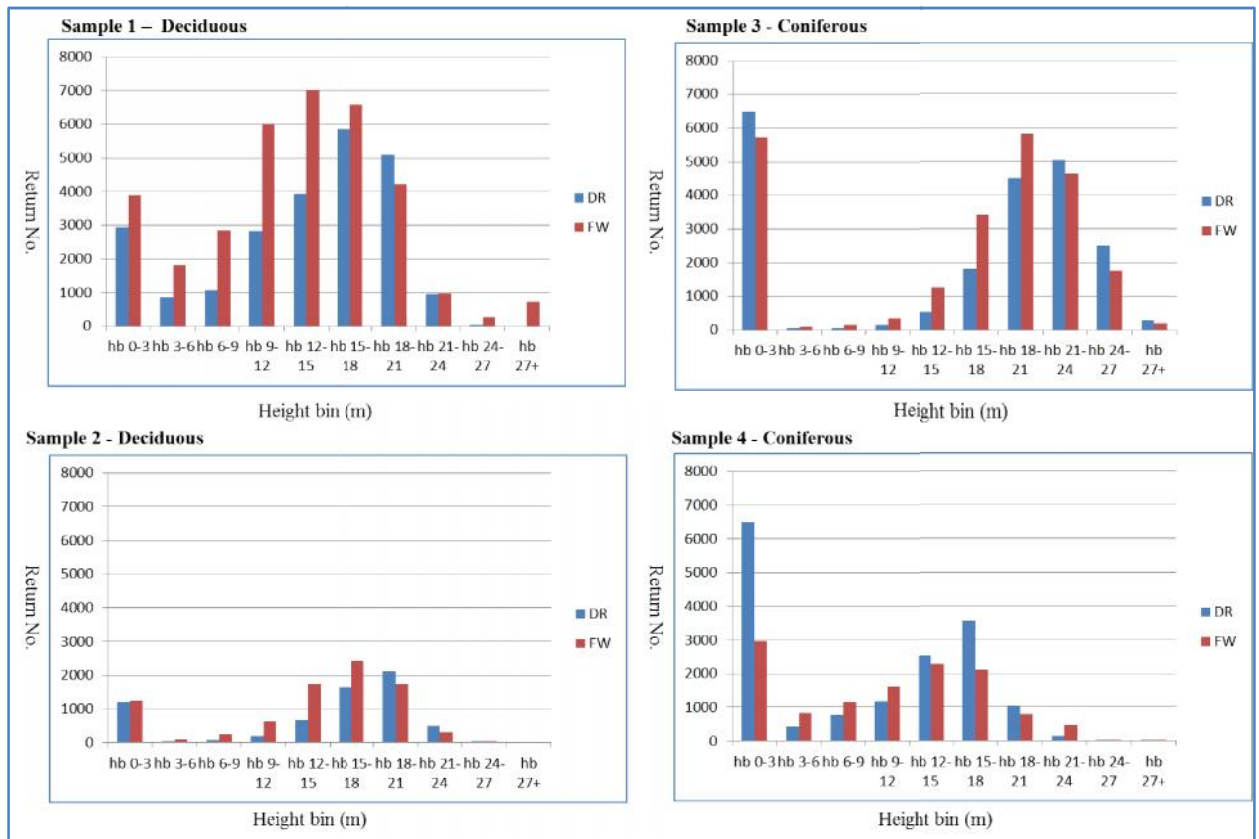


Figure 7.5 – Summary of height bin point counts for each of the sample areas in flight line 16 (both DR - blue and FW - red)

7.2 Estimating field-level forest structural and compositional metrics from DR LiDAR data

The statistical models applied to the DR LiDAR remote sensing dataset to estimate field plot-level metrics are presented here. Two statistical models, using Ordinary Least-Squares (OLS) and Akaike information criterion (AICc) multiple regression approaches, were calculated for each of the 39 field metrics using the inputs from the DR LiDAR remote sensing dataset. These inputs can be summarised as those statistics derived from ITC objects and area-based statistical summaries of LiDAR data layers. All ITC summary statistics were calculated using leaf-on LiDAR data. What follows is a breakdown of the regression models produced by forest structural, compositional, deadwood, regeneration and understorey metrics, and finally the two additional metrics required for index construction in further steps (see Chapter 9). In the models, the coefficients and variables were statistically significant at $p < 0.05$.

7.2.1 Forest structure within the plot

Table 7.3 gives the R^2 value for each of the regression models produced for DR LiDAR data for the estimation of the 14 structural metrics. When considering the statistical methods, OLS and AICc, the former performs better for eight of fourteen models, although eight of the R^2 values are very similar (± 0.1) to one another. The larger differences between the two methods are as follows: number of trees, mean DBH, standard deviation of DBH, average crown area, standard deviation of crown area and total crown area. AICc weights do not differ by much for the majority of the fourteen cases; typically the values vary between 1 and 2. Larger values are evident for the standard deviation of tree spacing and the standard deviation of the height of the first live branch

The input metrics for the 14 regression models are reported in Table 7.4. Many DR LiDAR metrics are used including from ITC-summaries, all and non-ground point metrics, percentiles, and intensity. ITC-summary metrics are common to almost all these models, followed by area-based summary metrics utilising all point distributions. Leaf-on metrics are used slightly more for model inputs than leaf-off metrics. When comparing OLS and AICc models for deriving average tree NN spacing, completely different model inputs are used, with the former having a higher R^2 . The OLS model uses one metric, the variance of all points in the leaf-on data, whereas the AICc model uses four inputs relating to the mean absolute deviation of points within the plot and three ITC-summary metrics for mean crown area, average NN spacing and number of overstorey trees.

Table 7.3 – Forest structural metrics, R^2 and AICc weights for DR LiDAR data (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	OLS R^2	AICc R^2	AICc weight	AICc of 2 nd best	AICc evidence ratio
<i>Number of trees per 30x30m</i>	<u>0.862</u>	0.672	0.140	0.087	1.609
<i>Average tree NN spacing (m)</i>	0.799	<u>0.906</u>	0.149	0.119	1.252
<i>STDev. Of tree NN spacing (m)</i>	<u>0.753</u>	0.682	0.835	0.165	5.061
<i>Mean DBH (cm)</i>	0.663	<u>0.796</u>	0.058	0.057	1.018
<i>STDev. of DBH (cm)</i>	0.625	<u>0.819</u>	0.056	0.041	1.366
<i>Basal area within 30x30m</i>	0.580	<u>0.658</u>	0.474	0.277	1.711
<i>Percentage big trees (DBH 40>x<80cm)</i>	0.628	<u>0.651</u>	0.031	0.031	1.000
<i>Percentage very big trees (DBH >80cm)</i>	<u>0.628</u>	0.471	0.134	0.116	1.155
<i>Percentage big trees (DBH >40cm)</i>	0.839	<u>0.920</u>	0.092	0.054	1.704
<i>Average height of the first live branch (m)</i>	<u>0.896</u>	0.876	0.057	0.038	1.500
<i>STDev. of the height of the first live branch (m)</i>	<u>0.824</u>	0.817	0.187	0.062	3.016
<i>Average crown horizontal area (m²)</i>	<u>0.748</u>	0.629	0.104	0.066	1.576
<i>STDev. crown horizontal area (m²)</i>	<u>0.857</u>	0.633	0.211	0.107	1.972
<i>Total crown horizontal area (m²)</i>	<u>0.748</u>	0.690	0.035	0.028	1.250

Table 7.4– DR LiDAR metric inputs for each forest structural metric regression model (Appendix C.1-C.3 should be consulted for a description of DR LiDAR metric names)

Metric Name:	LiDAR OLS model input metrics	LiDARAICc model input metrics
<i>Number of trees per 30x30m</i>	(1) TIF_m_area (ITC) (2) p020_2 (point-all) [leaf-on] (3) SHN_Med (point-non-ground) [leaf-off]	(1) SHA_Kur_2 (point-all) [leaf-on] (2) TIF_m_area (ITC) (3) TIF_TreeNo (ITC)
<i>Average tree spacing (m)</i>	(1) SHA_Var_2 (point-all) [leaf-on]	(1) SHA_MAD_2 (point-all) [leaf-on] (2) TIF_m_area (ITC) (3) TIF_Space (ITC) (4) TIF_TreeNo (ITC)
<i>STDev. Of tree spacing (m)</i>	(1) p045 (point-all) [leaf-off] (2) SHN_Kur (point-non-ground) [leaf-off]	(1) p060 (point-all) [leaf-off] (2) TIF_CR (ITC)
<i>Mean DBH (cm)</i>	(1) TIF_space (ITC) (2) TIF_meanCV (ITC)	(1) SHA_Var_2 (point-all) [leaf-on] (2) SIN_Kur (intensity-non-ground) [leaf-off] (3) TIF_m_area (ITC)
<i>STDev. of DBH (cm)</i>	(1) TIF_TreeNo (ITC) (2) p035_2 (point-all) [leaf-on] (3) SHN_Ske (point-non-ground) [leaf-off]	(1) p090i (Intensity-all) [leaf-off] (2) TIF_TreeNo (ITC)
<i>Basal area within 30x30m</i>	(1) p055i_2 (intensity-all) [leaf-on] (2) TIF_space (ITC) (3) SIA_Kur_2 (intensity-all) [leaf-on]	(1) p055i_2 (intensity-all) [leaf-on] (2) TIF_Space (ITC)
<i>Percentage big trees (DBH 40>x<80cm)</i>	(1) SHA_Var_2 (points-all) [leaf-on] (2) TIF_STD (ITC) (3) SHN_Var (points-non-ground) [leaf-off] (4) SHA_Kur (point-all) [leaf-off]	(1) SHA_Var_2 (point-all) [leaf-off] (2) TIF_space (ITC)
<i>Percentage very big trees (DBH >80cm)</i>	(1) TIF_CR (ITC)	(1) TIF_TreeNo (ITC)
<i>Percentage big trees (DBH >40cm)</i>	(1) SHA_STD_2 (points-all) [leaf-on] (2) TIF_m_area (ITC)	(1) SHA_STD_2 (points-all) [leaf-on] (2) TIF_Space (ITC) (3) TIF_TreeNo (points-all)
<i>Average height of the first live branch (m)</i>	(1) SHA_Var_2 (points-all) [leaf-on] (2) TIF_Space (ITC) (3) TIF_meanCV (ITC)	(1) lon_can_depth (points-all) [leaf-on] (2) SHA_STD (points-all) [leaf-off] (3) SIN_Ske (intensity-non-ground) [leaf-off]
<i>STDev. of the height of the first live branch (m)</i>	(1) SIA_Ske (intensity-all) [leaf-off] (2) TIF_Space (ITC) (3) p060i (intensity-all) [leaf-off] (4) SHN_Mean_2 (points-non-ground) [leaf-off]	(1) SHA_Max_2 (points-all) [leaf-on] (2) TIF_STD (ITC) (3) TIF_to_area (ITC)
<i>Average crown horizontal area (m²)</i>	(1) TIF_CR (ITC)	(1) SIN_Kur (intensity-non-ground) [leaf-off] (2) TIF_Space (ITC) (3) TIF_to_area (ITC)
<i>STDev. crown horizontal area (m²)</i>	(1) TIF_m_area (ITC) (2) SIA_Mean_2 (intensity-all) [leaf-on] (3) SIN_Kur (intensity-non-ground) [leaf-off]	(1) SIA_Med (intensity-all) [leaf-off] (2) TIF_m_area (ITC)
<i>Total crown horizontal area (m²)</i>	(1) SHA_Ske_2 (points-all) [leaf-on] (2) TIF_Space (ITC) (3) SHA_Ske (points-all) [leaf-off]	(1) SHA_Mean_2 (points-all) [leaf-on] (2) SHA_STD_2 (points-all) [leaf-on] (3) SIA_Var_2 (intensity-all) [leaf-on]

7.2.2 Forest composition within the plot

The R^2 values for each of the eight regression models for the prediction of forest compositional metrics is given in Table 7.5. Two of the field plot-level metrics could not be modelled using OLS regression, the SI (by stem count) and SH (by basal area). Only one metric, the number of tree species, had a lower R^2 value for the AICc models. The AICc evidence ratio indicated there was not much difference between the first and second ‘best’ models generated through the AICc approach.

DR LiDAR metric inputs into the regression models, as given in Table 7.6, included many types of area-based metrics for both point heights and intensity. Both leaf-on and leaf-off data were used. Intensity metrics saw greater usage for composition metrics than for structural metrics, as did the metric related to non-ground points.

Table 7.5 – Forest composition metrics, R^2 and AICc weights for DR LiDAR (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	OLS R^2	AICc R^2	AICc weight	AICc of 2 nd best	AICc evidence ratio
<i>Number of native trees per 30x30m</i>	0.428	<u>0.552</u>	0.133	0.093	1.430
<i>Percentage of native trees</i>	0.435	<u>0.528</u>	0.181	0.111	1.631
<i>No. of tree species per 30x30m</i>	<u>0.673</u>	0.426	0.104	0.104	1.000
<i>Shannon index (by stem count)</i>	0.571	<u>0.696</u>	0.126	0.095	1.326
<i>Simpson index (by stem count)</i>	-	0.589	0.055	0.043	1.279
<i>Shannon index (by BA)</i>	-	0.773	0.375	0.214	1.752
<i>Simpson index (by BA)</i>	0.448	<u>0.473</u>	0.035	0.034	1.029

Table 7.6 – DR LiDAR metric inputs for each forest compositional metric regression model (Appendix C.1-C.3 should be consulted for a description of DR LiDAR metric names)

Metric Name:	DR LiDAR OLS model input metrics	DR LiDAR AICc model input metrics
<i>Number of native trees per 30x30m</i>	(1) SHA_Kur_2 (points-all) [leaf-on] (2) p075_2 (points-all) [leaf-on]	(1) SHA_Ske (points-all) [leaf-off] (2) SHA_Ske_2 (points-all) [leaf-on] (3) TIF_m_area (ITC)
<i>Percentage of native trees</i>	(1) SIA_Kur (intensity-all) [leaf-off] (2) SIN_Var_2 (intensity-non-ground) [leaf-on] (3) p050 (points-all) [leaf-off]	(1) TIF_TreeNo (ITC)
<i>No. of tree species per 30x30m</i>	(1) SIN_STD_2 (intensity-non-ground) [leaf-on]	(1) p050 (points-all) [leaf-off] (2) SIN_Var_2 (intensity-non-ground) [leaf-on]
<i>Shannon index (by stem count)</i>	(1) SIN_Kur_2 (intensity-non-ground) [leaf-on] (2) p050i_2 (intensity-all) [leaf-on] (3) SHN_STD_2 (points-non-grounds) [leaf-on]	(1) SHN_Med (points-non-grounds) [leaf-off] (2) SIN_Ske_2 (intensity-non-ground) [leaf-on] (3) TIF_to_area (ITC) (4) TIF_totalCV (ITC)
<i>Simpson index (by stem count)</i>	-	(1) SHN_Mean_2 (points-non-grounds) [leaf-on] (2) SHN_STD_2 (points-non-grounds) [leaf-on] (3) SIA_Kur_2 (intensity-all) [leaf-on] (4) TIF_totalCV (ITC)
<i>Shannon index (by BA)</i>	-	(1) DEM_Slope_2 (points-ground) [leaf-on] (2) p015_2 (points-all) [leaf-on] (3) SHN_STD_2 (points-non-ground) [leaf-on] (4) SIA_Med (intensity-all) [leaf-off]
<i>Simpson index (by BA)</i>	(1) TIF_HT (ITC) (2) SIN_Ske_2 (intensity-non-ground)	(1) p070_2 (points-all) [leaf-on] (2) SIN_Ske_2 (intensity-non-grounds) [leaf-on]

7.2.3 Deadwood within the plot

The R^2 values for the regression models for the estimation of the four forest deadwood metrics are given in Table 7.7. Comparing the OLS and AICc models, three of the four of the latter had lower R^2 values, although the values for standing deadwood volume were very similar. AICc weights and evidence ratios indicate there was not much difference between the first and second ‘best’ models. DR LiDAR metric inputs into the regression models for deadwood included primarily metrics related to the LiDAR intensity of both ground and non-ground points, as seen in Table 7.8. Both leaf-on and leaf-off data were used. The standard deviation of ITC tree crown spacing was utilised for both OLS and AICc models for predicting the volume of standing deadwood.

Table 7.7 – Forest Deadwood metrics, R^2 and AICc weights for DR LiDAR data (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	OLS R^2	AICc R^2	AICc weight	AICc of 2 nd best	AICc evidence ratio
<i>Standing dead wood volume (m^3)</i>	<u>0.920</u>	0.906	0.113	0.105	1.076
<i>Standing dead wood decay class</i>	0.541	<u>0.883</u>	0.251	0.202	1.243
<i>Downed dead wood volume (m^3)</i>	<u>0.511</u>	0.420	0.024	0.019	1.263
<i>Downed dead wood decay class</i>	<u>0.786</u>	0.472	0.049	0.044	1.114

Table 7.8– DR LiDAR metric inputs for each forest Deadwood metric regression model (Appendix C.1-C.3 should be consulted for a description of DR LiDAR metric names)

Metric Name:	OLS model input metrics	AICc model input metrics
<i>Standing dead wood volume (m^3)</i>	(1) TIF_STD (ITC) (2) SIN_Kur (intensity-non-ground) [leaf-off] (3) SIG_Kur_2 (intensity-ground) [leaf-on]	(1) SIN_Kur (intensity-non-ground) [leaf-off] (2) SIN_Med (intensity-non-ground) [leaf-off] (3) TIF_STD (ITC)
<i>Standing dead wood decay class</i>	(1) SIA_Ske (intensity-all) [leaf-off] (2) SIG_Ske_2 (intensity-ground) [leaf-on]	(1) p015i_2 (intensity-all) [leaf-on] (2) SIN_Ske_2 (intensity-non-ground) [leaf-on] (2) SIN_STD (intensity-non-ground) [leaf-off]
<i>Downed dead wood volume (m^3)</i>	(1) SIG_MAD_2 (intensity-ground) [leaf-on]	(2) SHA_STD_2 (points-all) [leaf-on] (2) SIN_Ske (intensity-non-ground) [leaf-off]
<i>Downed dead wood decay class</i>	(1) SIN_Mean_2 (intensity-non-ground) [leaf-on]	(1) SIN_Med_2 (intensity-non-ground) [leaf-on]

7.2.4 Understorey metrics within the plot

There were 12 understorey field-level metrics in total; Table 7.9 gives the R^2 value for each regression model and AICc weights. Overall, R^2 values were high for both OLS and AICc models. Seven of the twelve regression model comparisons had very similar in the R^2 values, and five of the AICc models had lower R^2 values than the OLS counterpart. Three of these had a difference of up to ± 0.3 . One metric could not be estimated by OLS regression, the number of vascular species. R^2 values for the estimate of the SH index for native seedlings were low, for both OLS and AICc. Again, AICc weights and evidence ratio indicate there was not much difference between the first and second ‘best’ models.

The DR LiDAR metric inputs for each of the regression models are listed in Table 7.10. The majority of models utilise some measure of DR LiDAR return intensity as an input, whereas ITC-summary and point-height metrics are in the minority. AICc models typically use leaf-on data, where OLS generally includes both leaf-on and leaf-off data. There are a number of occurrences of intensity values from percentiles in the upper canopy being used in the models for estimating field-level metrics, these being the number of saplings and the number of seedling species. Measures of ground slope and roughness were used as inputs for estimating metrics on the forest floor.

Table 7.9 – Forest understorey metrics, R^2 and AICc weights for DR LiDAR data (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	OLS R^2	AICc R^2	AICc weight	AICc of 2nd best	AICc evidence ratio
<i>No. saplings per 30x30m</i>	<u>0.965</u>	0.922	0.152	0.135	1.126
<i>No. native saplings per 30x30m</i>	0.890	<u>0.896</u>	0.043	0.043	1.000
<i>Percentage of native saplings per 30x30m</i>	0.879	<u>0.951</u>	0.172	0.138	1.246
<i>No. of sapling species per 30x30m</i>	0.905	<u>0.938</u>	0.175	0.104	1.683
<i>No. seedlings per 30x30m</i>	<u>0.821</u>	0.527	0.073	0.050	1.460
<i>No. native seedlings per 30x30m</i>	0.413	<u>0.631</u>	0.026	0.022	1.182
<i>Percentage of native seedlings per 30x30m</i>	<u>0.920</u>	0.585	0.018	0.017	1.059
<i>No. of seedling species per 30x30m</i>	<u>0.739</u>	0.594	0.077	0.069	1.116
<i>Seedlings Shannon index for native species</i>	0.212	<u>0.212</u>	0.124	0.091	1.363
<i>No. vascular species</i>	-	<u>0.746</u>	0.054	0.044	1.227
<i>No. bryophyte species</i>	0.835	<u>0.858</u>	0.123	0.100	1.230
<i>Percentage of bare soil within 30x30m</i>	<u>0.761</u>	0.737	0.183	0.102	1.794

Table 7.10– DR LiDAR metric inputs for each forest understorey metric regression model (Appendix C.1-C.3 should be consulted for a description of DR LiDAR metric names)

Metric Name:	OLS model input metrics	AICc model input metrics
<i>No. saplings per 30x30m</i>	(1) p060i_2 (intensity-all) [leaf-on] (2) TIF_HT (ITC) (3) p045i_2 (intensity-all) [leaf-on] (4) SIN_Var (intensity-non-ground) [leaf-off]	(1) p060i_2 (intensity-all) [leaf-on] (2) p060_2 (points-all) [leaf-on] (3) SIA_Mean (intensity-all) [leaf-off]
<i>No. native saplings per 30x30m</i>	(1) p060i_2 (intensity-all) [leaf-on] (2) p050_2 (pointst-all) [leaf-on] (3) SIA_Mean (intensity-all) [leaf-off]	(1) p060i_2 (intensity-all) [leaf-on] (2) SHN_Med_2 (points-non-ground) [leaf-on]
<i>Percentage of native saplings per 30x30m</i>	(1) SHN_Ske_2 (points-non-ground) [leaf-on] (2) p095_2 (points-all) [leaf-on]	(1) SHN_Ske_2 (points-non-ground) [leaf-on] (2) SIN_Mean_2 (intensity-non-ground) [leaf-on] (3) SIN_Var (intensity-non-ground) [leaf-off]
<i>No. of sapling species per 30x30m</i>	(1) p045_2 (points-all) [leaf-on] (2) TIF_CR (ITC) (3) SIG_Med (intensity-ground) [leaf-off]	(1) p060i_2 (intensity-all) [leaf-on] (2) loff_can_depth (pointst_all) [leaf-off] (3) SIG_Med (intensity-ground) [leaf-off]
<i>No. seedlings per 30x30m</i>	(1) DEM_Slope (pointst-all) [leaf-off] (2) SIG_Kur (intensity-ground) [leaf-off] (3) SIN_Var (intensity-non-ground) [leaf-off] (4) p015i_2 (intensity-all) [leaf-on]	(1) SIG_MAD (intensity-ground) [leaf-off] (2) SIG_Mean_2 (intensity-ground) [leaf-on]
<i>No. native seedlings per 30x30m</i>	(1) TIF_STD (ITC) (2) TIF_meanCV (ITC)	(1) SIA_Mean (intensity-all) [leaf-off] (2) SIG_Kur (intensity-ground) [leaf-off] (3) SIG_Ske_2 (intensity-ground) [leaf-on] (4) TIF_meanCV (ITC)
<i>Percentage of native seedlings per 30x30m</i>	(1) SIG_Ske (intensity-ground) [leaf-off] (2) SIG_Kur (intensity-ground) [leaf-off] (3) SIN_Mean (intensity-non-ground) [leaf-off] (4) SHN_Kur_2 (points-non-ground) [leaf-on]	(1) p045_2 (points-all) [leaf-on] (2) SIA_Med_2 (intensity-all) [leaf-on]
<i>No. of seedling species per 30x30m</i>	(1) p060i_2 (intensity-all) [leaf-on] (2) p060_2 (points-all) [leaf-on] (3) CR_ratio (points-all) [leaf-off]	(1) DEM_Slope_2 (points-ground) [leaf-on] (2) SHA_Mean_2 (points-all) [leaf-on] (3) SIA_Ske_2 (intensity-all) [leaf-on]
<i>Seedlings Shannon Index for native species</i>	(1) TIF_m_area (ITC)	(1) SIG_Ske_2 (intensity-ground) [leaf-on]
<i>No. Vascular Species</i>	-	(1) DEM_Rough (points-ground)[leaf-off] (2) p025i_2 (intensity-all) [leaf-on] (3) SHA_Ske_2 (points-all) [leaf-on] (4) TIF_HT (ITC)
<i>No. Bryophyte Species</i>	(1) p085i (intensity-all) [leaf-off] (2) lon_max_gap (points-all) [leaf-on] (3) TIF_m_area (ITC) (4) TIF_Space (ITC)	(1) lon_can_layer (points-all) [leaf-on] (2) p010_2 (points-all) [leaf-on] (3) SIN_Kur (intensity-non-ground) [leaf-off] (4) SIN_STD (intensity-non-ground) [leaf-off] (5) TIF_STD (ITC)
<i>Percentage of bare soil within 30x30m</i>	(1) p020_2 (points-all) [leaf-on] (2) SHA_Var (pointst-all) [leaf-off]	(1) p025i_2 (intensity-all) [leaf-on]

7.2.5 Metrics required for condition index construction

Regression modelling techniques were applied to the two field data derived metrics required for input into two condition indices, the R^2 values and AICc weights are given for the models in Table 7.11. The R^2 values for the sum of the DBH differences are high and similar between OLS and AICc. LiDAR intensity and ITC-summary metrics are both used, and in addition, the OLS method also uses a metric relating to the distribution of point heights. The regression models for the index of vertical separation differ, and the OLS model has the lower R^2 value. The regression models are listed in Table 7.12. The input metrics for the index of vertical separation differ between the two models; OLS uses non ground point distributions, whilst AICc utilises metrics related to the crown shape and coverage. AICc weights and evidence ratio indicate there was a slight difference between the first and second ‘best’ models generated through the AICc approach.

Table 7.11 – Condition index construction metrics, R^2 and AICc weights for DR LiDAR (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	OLS R^2	AICc R^2	AICc weight	AICc of 2 nd best	AICc evidence ratio
<i>Sum of the DBH differences between neighbouring trees (cm)</i>	0.873	<u>0.924</u>	0.155	0.092	1.685
<i>Index of Vertical Separation</i>	0.526	<u>0.743</u>	0.152	0.093	1.634

Table 7.12 – DR LiDAR metric inputs for each forest condition index metric regression model (Appendix C.1-C.3 should be consulted for a description of DR LiDAR metric names.)

Metric Name:	OLS model input metrics	AICc model input metrics
<i>Sum of the DBH differences between neighbouring trees (cm)</i>	(1) TIF_TreeNo (ITC) (2) SHA_Kur_2 (points-all) [leaf-on] (3) p045i_2 (intensity-all) [leaf-on]	(1) p015i_2 (intensity-all) [leaf-on] (2) SIA_MAD_2 (intensity-all) [leaf-on] (3) TIF_m_area (ITC) (4) TIF_meanCV (ITC)
<i>Index of Vertical Separation</i>	(1) SHN_Ske (points-non-ground) [leaf-off] (2) SHN_STD_2 (points-non-ground) [leaf-on]	(1) CC (points-all) [leaf-off] (2) CR_ratio (pointst-all) [leaf-off] (3) TIF_to_area (ITC)

7.3 Validation of DR LiDAR remote sensing derived model estimates against field data

The various direct and statistical model estimates of forest structure and composition are compared with fieldwork data within this section. It begins with the direct comparisons which can be made using DR LIDAR remote sensing data, followed by a summary of the various statistical model estimates derived from remote sensing. RMSE and NRMSE values are presented to illustrate the accuracy of each model prediction. NRMSE is a modified RMSE value to account for the field data population range, and produces a value of 0 to 1. NRMSE values of 0.5 or greater indicate poor model performance.

7.3.1 Direct comparisons between field measurements and remote sensing metrics

Direct comparisons involved: (i) the number of trees within the plot extent, (ii) average tree height, (iii) number of canopy layers, (iv) canopy openness, (v) individual stem distances and (vi) tree crown horizontal diameters.

Table 7.13 summarises the relationships between the automated tree detection techniques and the actual stem counts within the plot, and those stems identified as overstorey (defined as trees with a height >15m). Automated individual tree crown (ITC) delineation techniques were conducted in the two pieces of software, eCognition and TIFFS, each of which were employed to delineate individual tree crowns using leaf-on DR LiDAR data only. The eCognition approach was mentioned in chapter 6, where it was utilised to classify tree species ITC objects, but without a direct comparison with stem count.

Both of the automated approaches underestimate the amount of stems within the plot, with the one exception of plot 33. This underestimation is especially true of very high tree density plots, such as plot 23 and 30. If one considers the stem count of overstorey trees only, the automated approaches more closely resemble the field data; this is also reflected in the lower RMSE and NRMSE values for overstorey trees.

Table 7.13 – Comparison of stem counts between field data and automated tree detection

Site no.	Field total stem no.	Over- storey stem no.	TIFFS ITC tree no.	eCognition TC tree no.
22	20	15	11	10
23	71	54	25	33
24	53	30	15	14
25	21	15	12	10
26	17	13	8	10
27	33	13	9	11
28	35	34	13	21
29	30	21	11	12
30	90	80	21	22
31	24	24	27	18
32	20	17	17	19
33	28	27	30	36
34	28	17	16	20
35	33	31	19	23
36	25	22	18	22
37	23	16	12	11
38	36	28	16	16
39	16	15	13	23
40	17	13	13	12
41	24	14	15	18
RMSE AgainstField total stem no.				
		TIFFS tree no.		eCognition tree no.
RMSE		23.48		22.13
NRMSE		0.32		0.30
RMSE Against Over-storey stem no.				
		TIFFS tree no.		eCognition tree no.
RMSE		16.58		15.52
NRMSE		0.25		0.23

Area-based estimates of canopy structure are summarised in Table 7.14. The first of these, Lorey's height, is an average estimate of tree heights weighted by basal area. Average height was estimated by an average of ITC objects generated by the TIFFS software using leaf-on DR LiDAR. The RMSE and NRMSE values for the TIFFS average height were 3.46m and 0.18 respectively, and indicate a potential maximum ± 3.26 m deviation between field and remote sensing measurements. The total number of canopy layers relates to the height layers of vegetation within a plot, and typically varied between one and three for the validation field plots. LiDAR estimates were made using the SPDlib software for leaf-on DR LiDAR data, which produced an RMSE of 1.18. With two exceptions, the majority of layer estimates were correct or within ± 1 of the field counterpart.

Canopy openness was estimated in the field using aggregated data from hemispherical photographs. The comparison between field and DR LiDAR estimates are shown in Table 7.15. RMSE values were calculated as 25.59% and 13.29% for leaf-off and leaf-on datasets respectively.

Table 7.14 – Comparison of forest structural metrics from DR LiDAR data with field data

Site no.	Lorey's Height (m) [Field]	TIFFS average height (m) [DR]	No. Canopy Layers [Field]	No. canopy layers [DR]
22	12.83	16.91	2	3
23	16.85	11.97	2	2
24	19.17	21.42	2	3
25	22.13	24.12	2	3
26	21.05	20.38	1	3
27	24.79	30.55	3	5
28	23.35	20.36	1	3
29	17.89	22.56	2	4
30	19.38	20.58	2	3
31	28.78	29.10	1	2
32	25.64	21.10	1	3
33	28.36	26.42	1	2
34	22.99	24.20	3	3
35	23.89	26.81	3	3
36	25.12	23.71	2	2
37	17.91	22.48	2	3
38	21.4	17.64	3	3
39	33.12	26.77	3	3
40	24.45	21.82	3	3
41	18.99	21.21	2	3
	RMSE	3.46	RMSE	1.18
	NRMSE	0.18	NRMSE	0.39

Table 7.15 – Canopy openness comparison of DR LiDAR data with field data

Site no.	Canopy openness (%) [Field]	Canopy openness (%) (Leaf off) [DR]	Canopy openness (%) (Leaf on) [DR]
22	20.59	44.20	32.40
23	17.24	40.80	33.90
24	19.59	46.80	38.80
25	21.01	44.80	26.90
26	22.81	39.80	33.90
27	12.97	40.20	18.10
28	26.84	54.00	49.50
29	22.76	36.80	35.00
30	15.68	44.90	20.20
31	14.67	46.60	16.30
32	15.8	41.60	18.80
33	18.89	40.20	34.30
34	15.72	51.20	32.30
35	21.33	47.80	43.60
36	14.75	40.50	6.20
37	23.87	46.00	37.90
38	17.72	41.70	36.60
39	22.8	54.10	36.40
40	16.21	44.50	13.00
41	10.35	26.20	16.60
	RMSE	25.59	13.29

7.3.2 Validation of remote sensing derived statistical models against field data

This section contains the results of the regression models derived from DR LIDAR datasets compared to 2012 field data measurements. RMSEs and NRMSEs were calculated for each of the remote sensing derived models. An overview of the OLS and AICc models for each field-level metric is presented in Table 7.16 for overstorey metrics and in Table 7.17 for understorey metrics. Table 7.16 includes a number of ITC summary values for the number of trees within an area, NN spacing, average crown and total crown areas.

A number of metrics had poor RMSE and NRMSE values for both regression methods: percentage very big trees (DBH >80cm), SH index (by basal area), percentage of native saplings, number of sapling species and the number of bryophyte species. As before, poor values are defined arbitrarily as NRMSE values above 0.5, and in consideration of the range of actual field values. There are two examples of one regression method producing NRSME values above 0.5 (for the percentage of native trees, and percentage of bare soil), whilst the other does not. In relation to the statistically derived estimates for the standard deviation of

NN spacing, very small ranges were encountered for this metric in the field, where RMSE of OLS corresponds to less than 1m.

Seventeen out of the remaining thirty-four models showed OLS methods produced better results than AICc, although some were similar. For example models for estimating total basal area had RMSE values of 0.848m² and 0.825m² for OLS and AICc respectively. In reference to Table 7.16, ITC-summary metrics produced similar RMSE values for all field metric estimations, apart from those related to stem NN spacing. Average crown and total crown area ITC estimates produced results similar or slightly better than those derived statistically.

Table 7.16 – The OLS and AICc DR LiDAR regression models. Part 1: overstorey composition.

Field metric	ITC Direct RMSE*	ITC Direct NRMSE*	OLS RMSE	OLS NRMSE	AICc RMSE	AICc NRMSE
<i>Number of trees per 30x30m</i>	24.077	0.325	18.544	0.251	<u>18.035</u>	<u>0.244</u>
<i>Number of native trees per 30x30m</i>	-	-	15.464	0.297	<u>15.045</u>	<u>0.289</u>
<i>Percentage of native trees</i>	-	-	73.622	0.736	<u>28.753</u>	<u>0.288</u>
<i>No. of tree species per 30x30m</i>	-	-	1.628	0.407	<u>1.500</u>	<u>0.375</u>
<i>Average tree NN spacing (m)</i>	4.234	1.275	<u>1.097</u>	<u>0.330</u>	1.842	0.555
<i>STDev. Of tree NN spacing (m)</i>	0.657	0.790	0.577	0.695	0.958	1.153
<i>Mean DBH (cm)</i>	-	-	<u>8.357</u>	<u>0.287</u>	11.925	0.409
<i>STDev. of DBH (cm)</i>	-	-	<u>10.572</u>	<u>0.217</u>	11.922	0.245
<i>Basal area within 30x30m (m²)</i>	-	-	0.848	0.264	<u>0.825</u>	<u>0.257</u>
<i>(i) Percentage big trees (DBH 40>x<80cm)</i>	-	-	<u>20.028</u>	<u>0.286</u>	24.762	0.354
<i>(ii) Percentage very big trees (DBH >80cm)</i>	-	-	9.845	1.624	7.042	1.162
<i>(iii) Percentage big trees (DBH >40cm)</i>	-	-	<u>26.503</u>	<u>0.379</u>	27.052	0.386
<i>Shannon index (by stem count)</i>	-	-	<u>0.526</u>	<u>0.367</u>	0.625	0.436
<i>Simpson index (by stem count)</i>	-	-	-	-	0.306	0.237
<i>Shannon index (by BA)</i>	-	-	-	-	0.546	0.711
<i>Simpson index (by BA)</i>	-	-	0.662	0.229	<u>0.649</u>	<u>0.225</u>
<i>Average height of the first live branch (m)</i>	-	-	3.212	0.296	<u>2.679</u>	<u>0.247</u>
<i>STDev. of the height of the first live branch (m)</i>	-	-	2.309	0.448	<u>1.660</u>	<u>0.322</u>
<i>Average crown horizontal area (m²)</i>	25.459	0.292	36.700	0.421	<u>23.532</u>	<u>0.270</u>
<i>STDev. crown horizontal area (m²)</i>	-	-	<u>20.623</u>	<u>0.361</u>	26.788	0.468
<i>Total crown horizontal area (m²)</i>	558.118	0.387	<u>561.549</u>	<u>0.389</u>	641.159	0.444

*Denotes this was not derived statistically and cannot be applied to all metrics

Table 7.17– The OLS and AICc DR LiDAR regression models. Part 2: understorey composition.

Field metric	OLS RMSE	OLS NRMSE	AICc RMSE	AICc NRMSE
<i>Standing dead wood volume (m³)</i>	<u>1.497</u>	<u>0.181</u>	2.927	0.354
<i>Standing dead wood decay class</i>	0.246	0.369	<u>0.243</u>	<u>0.364</u>
<i>Downed dead wood volume (m³)</i>	<u>2.737</u>	<u>0.301</u>	2.862	0.315
<i>Downed dead wood decay class</i>	<u>0.170</u>	<u>0.362</u>	0.174	0.371
<i>No. saplings per 30x30m</i>	26.039	0.241	<u>24.406</u>	<u>0.226</u>
<i>No. native saplings per 30x30m</i>	<u>10.402</u>	<u>0.274</u>	12.350	0.325
<i>Percentage of native saplings per 30x30m</i>	73.948	0.739	60.853	0.609
<i>No. of sapling species per 30x30m</i>	2.012	0.671	1.993	0.664
<i>No. seedlings per 30x30m</i>	<u>295.244</u>	<u>0.315</u>	337.836	0.361
<i>No. native seedlings per 30x30m</i>	<u>239.082</u>	<u>0.332</u>	243.125	0.338
<i>Percentage of native seedlings per 30x30m</i>	49.305	0.493	<u>44.049</u>	<u>0.440</u>
<i>No. of seedling species per 30x30m</i>	<u>3.162</u>	<u>0.243</u>	3.465	0.267
<i>Seedlings Shannon index for native species</i>	1.073	0.665	<u>0.726</u>	<u>0.450</u>
<i>No. vascular species</i>	<u>2.356</u>	<u>0.393</u>	3.455	0.576
<i>No. bryophyte species</i>	9.296	0.930	7.795	0.780
<i>Percentage of bare soil within 30x30m</i>	50.690	0.534	<u>44.041</u>	<u>0.464</u>
<i>Sum of the DBH differences between neighbouring trees (cm)</i>	2.712	0.091	<u>2.099</u>	<u>0.070</u>
<i>Index of vertical separation</i>	<u>0.161</u>	<u>0.319</u>	1.029	2.038

7.4 Estimating field-level forest structural and compositional metrics from FW LiDAR data

A listing of the statistical models applied to the FW LiDAR remote sensing datasets to estimate field plot-level metrics is presented here. Each of the models presented were statistically significant ($p < 0.05$).

The two regression approaches, OLS and AICc, were applied to the FW LiDAR data in order to predict the 39 field metrics. These inputs can be summarised as those statistics derived from ITC objects and area-based statistical summaries of FW LiDAR data layers. What follows is a breakdown of the regression models produced by forest structural, compositional, deadwood, regeneration and understorey metrics, finally, the two additional metrics required for index construction.

7.4.1 Forest structure within the plot

There are 14 forest structural metrics estimated through the application of regression models from FW LiDAR data, these are listed in Table 7.18 which gives their associated R^2 values. Comparing R^2 values for OLS and AICc, it is OLS which is consistently higher, for all but three models, the average tree NN spacing, the standard deviation of height to the first live branch and the total crown area. However, many of the R^2 values were similar, and varied at most by 0.2.

The difference between first and second ‘best’ AICc weights was low for all but two of the field metric regression models. The standard deviation of tree spacing evidence ratio showed that the first model was over five times better than the second model. The evidence ratio of the standard deviation of height to the first live branch showed the first model was over three times better.

The FW metric inputs to the regression equations are given in Table 7.19. These metrics include ITC-summaries, point-heights, curve-amplitude and curve-width. An array of leaf-on and leaf-off metrics was used. The most common metrics were point-heights and ITC-summary metrics. Metrics related to DBH and basal area utilise both amplitude and width metrics.

Table 7.18 – Forest structure metrics, R^2 and AICc weights for FW LiDAR data (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	OLS R^2	AICc R^2	AICc weight	AICc of 2nd best	AICc evidence ratio
<i>Number of trees per 30x30m</i>	<u>0.897</u>	0.672	0.140	0.087	1.609
<i>Average tree NN spacing (m)</i>	0.786	<u>0.906</u>	0.149	0.119	1.252
<i>STDev. of tree NN spacing (m)</i>	<u>0.849</u>	0.682	0.835	0.165	5.061
<i>Mean DBH (cm)</i>	<u>0.933</u>	0.796	0.058	0.057	1.018
<i>STDev. of DBH (cm)</i>	<u>0.823</u>	0.819	0.056	0.041	1.366
<i>Basal area within 30x30m</i>	<u>0.691</u>	0.658	0.474	0.277	1.711
<i>Percentage big trees (DBH 40>x<80cm)</i>	<u>0.913</u>	0.651	0.031	0.031	1.000
<i>Percentage very big trees (DBH >80cm)</i>	<u>0.642</u>	0.471	0.134	0.116	1.155
<i>Percentage big trees (DBH >40cm)</i>	<u>0.939</u>	0.920	0.092	0.054	1.704
<i>Average height of the first live branch (m)</i>	<u>0.947</u>	0.876	0.057	0.038	1.500
<i>STDev. of the height of the first live branch (m)</i>	0.509	<u>0.817</u>	0.187	0.062	3.016
<i>Average crown horizontal area (m²)</i>	<u>0.862</u>	0.629	0.104	0.066	1.576
<i>STDev. crown horizontal area (m²)</i>	<u>0.811</u>	0.633	0.211	0.107	1.972
<i>Total crown horizontal area (m²)</i>	0.521	<u>0.690</u>	0.035	0.028	1.250

Table 7.19 – FW LiDAR metric inputs for each forest structural metric regression model (Appendix C.3, C.6-C.8 should be consulted for a description of FW LiDAR metric names)

Metric Name:	OLS model input metrics	AICc model input metrics
<i>Number of trees per 30x30m</i>	(1) TIF_m_area (ITC) (2) ht_p20_2 (points-all) [leaf-on] (3) kurngamp (amplitude-non-ground) [leaf-off] (4) skwnght (points-non-ground) [leaf-off]	(1) kurht_2 (points-all) [leaf-on] (2) TIF_TreeNo (ITC)
<i>Average tree spacing (m)</i>	(1) TIF_Space (ITC) (2) amp_p20 (amplitude-all) [leaf-off]	(1) TIF_Space (ITC) (2) TIF_to_area (ITC) (3) varht_2 (points-all) [leaf-on] (4) wd_p35 (width-all) [leaf-off]
<i>STDev. of tree spacing (m)</i>	(1) TIF_m_area (ITC) (2) meanht_2 (points-all) [leaf-on] (3) kurwid (width-all) [leaf-off]	(1) ht_p40 (points-all) [leaf-off] (2) meanwid (width-all) [leaf-off]
<i>Mean DBH (cm)</i>	(1) TIF_m_area (ITC) (2) varht_2 (points-all) [leaf-on] (3) ht_p30 (points-all) [leaf-off] (4) kurngamp (amplitude-non-ground) [leaf-off] (5) wd_p15_2 (width-all) [leaf-on]	(1) TIF_m_area (ITC) (2) varngwid_2 (width-non-ground) [leaf-on]
<i>STDev. of DBH (cm)</i>	(1) kurngwid (width-non-ground) [leaf-off] (2) amp_p15 (amplitude-all) [leaf-off] (3) wd_p25_2 (width-all) [leaf-on]	(1) amp_p65 (amplitude-all) [leaf-off] (2) TIF_m_area (ITC) (3) wd_p25_2 (width-all) [leaf-on]
<i>Basal area within 30x30m</i>	(1) kurngamp (amplitude-non-ground) [leaf-off] (2) TIF_Space (ITC) (3) TIF_HT (ITC) (4) wd_p25 (width-all) [leaf-off]	(1) ht_p60_2 (points-all) [leaf-on] (2) kurngamp (amplitude-non-ground) [leaf-off] (3) TIF_to_area (ITC) (4) varamp_2 (amplitude-all) [leaf-on]
<i>Percentage big trees (DBH 40>x<80cm)</i>	(1) varht (points-all) [leaf-off] (2) TIF_Space (ITC) (3) ht_p30 (points-all) [leaf-off]	(1) ht_p70 (points-all) [leaf-off] (2) varnght (points-non-ground) [leaf-off]
<i>Percentage very big trees (DBH >80cm)</i>	(1) TIF_m_area (ITC)	(1) stdngwid (width-non-ground) [leaf-off]
<i>Percentage big trees (DBH >40cm)</i>	(1) varht_2 (points-all) [leaf-on] (2) TIF_Space (ITC) (3) kurwid (width-all) [leaf-off]	(1) adevht_2 (points-all) [leaf-on] (2) TIF_Space (ITC)
<i>Average height of the first live branch (m)</i>	(1) loff_max_gap (points-all) [leaf-off] (2) TIF_Space (ITC) (3) TIF_totalCV (ITC) (4) varngamp_2 (amplitude-non-ground) [leaf-on]	(1) loff_max_gap (points-all) [leaf-off] (2) TIF_Space (ITC)
<i>STDev. of the height of the first live branch (m)</i>	(1) TIF_m_area (ITC) (2) TIF_meanCV (ITC)	(1) lon_can_depth (points-all) [leaf-on] (2) maxht (points-all) [leaf-off] (3) TIF_meanCV (ITC) (4) TIF_to_area (ITC)
<i>Average crown horizontal area (m²)</i>	(1) TIF_m_area (ITC) (2) kurngamp (amplitude-non-ground) [leaf-off] (2) amp_p15 (amplitude-all) [leaf-off]	(1) adevngwid (width-non-ground) [leaf-off] (2) amp_p50_2 (amplitude-all) [leaf-on] (3) TIF_meanCV (ITC)
<i>STDev. crown horizontal area (m²)</i>	(1) TIF_m_area (ITC) (1) medamp_2 (amplitude-all) [leaf-on]	(1) kurwid (width-all) [leaf-off] (2) meanamp (amplitude-all) [leaf-off]
<i>Total crown horizontal area (m²)</i>	(1) ht_p30_2 (points-all) [leaf-on] (2) TIF_Space (ITC)	(1) amp_p35_2 (amplitude-all) [leaf-on] (2) meanht_2 (points-all) [leaf-on] (3) TIF_Space (ITC)

7.4.2 Forest composition within the plot

A listing of the seven forest compositional metrics and their associated R^2 values are given in Table 7.20. Again, OLS regression models produce higher R^2 values than their AICc counterparts for all but one metric, the number of native trees. R^2 values are similar however. The AICc evidence ratio indicated there was not much difference between the first and second ‘best’ models.

FW LiDAR metrics used for model inputs (see Table 7.21) are varied, but both curve-amplitude and curve-width metrics see an increased usage in relation to the identification of species diversity. Both leaf-on and leaf-off datasets were used.

Table 7.20 – Forest composition metrics, R^2 and AICc weights for FW LiDAR data (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	OLS R^2	AICc R^2	AICc weight	AICc of 2 nd best	AICc evidence ratio
<i>Number of native trees per 30x30m</i>	0.460	<u>0.614</u>	0.102	0.057	1.789
<i>Percentage of native trees</i>	<u>0.600</u>	0.480	0.081	0.057	1.421
<i>No. of tree species per 30x30m</i>	<u>0.576</u>	0.320	0.114	0.070	1.629
<i>Shannon index (by stem count)</i>	<u>0.672</u>	0.627	0.072	0.068	1.059
<i>Simpson index (by stem count)</i>	<u>0.546</u>	0.488	0.072	0.060	1.200
<i>Shannon index (by BA)</i>	<u>0.718</u>	0.592	0.062	0.050	1.240
<i>Simpson index (by BA)</i>	<u>0.612</u>	0.570	0.012	0.012	1.000

Table 7.21 – FW LiDAR metric inputs for each forest compositional metric regression model (Appendix C.3, C.6-C.8 should be consulted for a description of FW LiDAR metric names)

Metric Name:	OLS model input metrics	AICc model input metrics
<i>Number of native trees per 30x30m</i>	(1) wd_p30_2 (width-all) [leaf-on] (2) ht_p45 (points-all) [leaf-off]	(1) kurht_2 (points-all) [leaf-on]
<i>Percentage of native trees</i>	(1) TIF_TreeNo (ITC) (2) kurngwid_2 (width-non-ground) [leaf-on]	(2) TIF_TreeNo (ITC)
<i>No. of tree species per 30x30m</i>	(1) wd_p30_2 (width-all) [leaf-on] (2) TIF_to_area (ITC) (3) skwngamp_2 (amplitude-non-ground) [leaf-on]	(1) skwngamp_2 (amplitude-non-ground) [leaf-on] (2) wd_p20_2 (width-all) [leaf-on]
<i>Shannon index (by stem count)</i>	(1) kurngamp_2 (amplitude_non-ground) [leaf-on] (2) TIF_to_area (ITC) (3) mednght (points-non-ground) [leaf-off]	(1) skwngamp_2 (amplitude-non-ground) [leaf-on] (2) TIF_to_area (ITC) (3) varht_2 (points-all) [leaf-on]
<i>Simpson index (by stem count)</i>	(1) kurngamp_2 (amplitude-non-ground) [leaf-on] (2) ht_p45_2 (points-all) [leaf-on] (3) ht_p40 (points-all) [leaf-off]	(1) adevwid (width-all) [leaf-off] (2) stdht_2 (points-all) [leaf-on] (3) wd_p45_2 (width-all) [leaf-on]
<i>Shannon index (by BA)</i>	(1) ht_p15_2 (point-all) (2) TIF_m_area (ITC) (3) wd_p15_2 (width-all) (4) wd_p35 (width-all)	(1) maxgwid (width-ground) (2) stdnght_2 (points-non-ground) (3) varamp_2 (amplitude-all)
<i>Simpson index (by BA)</i>	(1) kurht (points-all) (2) stdnght_2 (points-non-ground) (3) skwngamp_2 (amplitude-non-ground)	(1) medamp_2 (amplitude-all) (2) stdht_2 (points-all)

7.4.3 Deadwood within the plot

There was a total of four deadwood metrics to be predicted through regression models, the R^2 values and AICc weights are presented in Table 7.22. Again R^2 values for OLS and AICc models are similar although OLS R^2 values are higher for all but one metric, the standing deadwood decay class. AICc weights and evidence ratios indicate there was not much difference between the first and second ‘best’ models. The most common attributes used in the regression models are related to LiDAR amplitude, closely followed by curve-width. Table 7.23 gives the FW LiDAR input metrics used in the regression models. Both models for determining standing deadwood volume use the same inputs, albeit in a different order. Inputs for DDW volume share the same first input. This may explain the similar R^2 values. Again, a mix of leaf-on and leaf-off datasets was used.

Table 7.22 – Forest deadwood metrics, R^2 and AICc weights for FW LiDAR data (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	OLS R^2	AICc R^2	AICc weight	AICc of 2 nd best	AICc evidence ratio
<i>Standing dead wood volume (m^3)</i>	<u>0.948</u>	0.919	0.080	0.078	1.026
<i>Standing dead wood decay class</i>	0.470	<u>0.589</u>	0.033	0.032	1.031
<i>Downed dead wood volume (m^3)</i>	<u>0.451</u>	0.431	0.047	0.041	1.146
<i>Downed dead wood decay class</i>	<u>0.749</u>	0.569	0.078	0.064	1.219

Table 7.23 – FW LiDAR metric inputs for each forest deadwood metric regression model (Appendix C.3, C.6-C.8 should be consulted for a description of FW LiDAR metric names)

Metric Name:	FW LiDAR OLS model input metrics	FW LiDAR AICc model input metrics
<i>Standing dead wood volume (m^3)</i>	(1) TIF_STD (ITC) (2) wd_p25_2 (width-all) [leaf-on] (3) skwngamp (amplitude-non-ground) [leaf-on]	(1) skwngamp (amplitude-non-ground) [leaf-off] (2) TIF_STD (ITC) (3) wd_p25_2 (width-all) [leaf-on]
<i>Standing dead wood decay class</i>	(1) maxgwid (width-ground) [leaf-off]	(1) amp_p35_2 (amplitude-all) [leaf-on] (2) ht_p20_2 (points-all) [leaf-on]
<i>Downed dead wood volume (m^3)</i>	(1) skwngamp (amplitude-non-ground) [leaf-off] (2) adevht_2 (points-all) [leaf-on]	(1) skwngamp (amplitude-non-ground) [leaf-off] (2) wd_p25_2 (width-all) [leaf-on]
<i>Downed dead wood decay class</i>	(1) varngamp_2 (amplitude-non-ground) [leaf-on] (2) lon_can_layer (points-all) [leaf-on] (3) skwgwid_2 (width-ground) [leaf-on]	(1) meanngamp_2 (amplitude-non-ground) [leaf-on] (2) wd_p60_2 (width-all) [leaf-on]

7.4.4 Understorey metrics within the plot

There were 12 forest understorey structural and composition metrics in total, Table 7.24 contains a listing of the regression model R^2 values and AICc weights for each of these metrics. Neither regression method was effective at modelling the seedlings Shannon Index for native species. OLS could not model the number of saplings or the number of native saplings. Overall, AICc models had higher R^2 values for all but four of the remaining models, (and for three of these, AICc R^2 values were within 0.2 of those of OLS). AICc weights and evidence ratios indicate there was not much difference between the first and second ‘best’ models.

Each of the FW LiDAR metrics used as a predictor variable in the OLS and AICc regression models are reported in Table 7.25. Field metrics related to saplings are correlated with the FW metrics concerning point-heights, curve-width and amplitude for all points. Seedlings correlate mainly with ground metrics for curve-amplitude and width. Number of vascular and bryophyte species correlate with non-ground and all metrics for point-heights and amplitude. Percentage bare soil is related to point-heights and ITC only. Both leaf-on and leaf-off metrics were used.

Table 7.24 – Forest understorey metrics, R^2 and AICc weights for FW LiDAR data (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	OLS R^2	AICc R^2	AICc weight	AICc of 2 nd best	AICc evidence ratio
<i>No. saplings per 30x30m</i>	-	0.973	0.448	0.442	1.014
<i>No. native saplings per 30x30m</i>	-	0.968	0.503	0.497	1.012
<i>Percentage of native saplings per 30x30m</i>	0.803	<u>0.834</u>	0.256	0.193	1.326
<i>No. of sapling species per 30x30m</i>	0.702	<u>0.954</u>	0.567	0.299	1.896
<i>No. seedlings per 30x30m</i>	0.282	<u>0.468</u>	0.055	0.049	1.122
<i>No. native seedlings per 30x30m</i>	<u>0.682</u>	0.535	0.124	0.083	1.494
<i>Percentage of native seedlings per 30x30m</i>	<u>0.683</u>	0.432	0.050	0.045	1.111
<i>No. of seedling species per 30x30m</i>	0.206	<u>0.447</u>	0.113	0.064	1.766
<i>Seedlings Shannon index for native species</i>	-	-	-	-	-
<i>No. vascular species</i>	0.194	0.194	0.109	0.077	1.416
<i>No. bryophyte species</i>	<u>0.782</u>	0.457	0.269	0.188	1.431
<i>Percentage of bare soil within 30x30m</i>	<u>0.862</u>	0.749	0.083	0.069	1.203

Table 7.25 – FW LiDAR metric inputs for each forest understorey metric regression model (Appendix C.3, C.6-C.8 should be consulted for a description of FW LiDAR metric names)

Metric Name:	OLS model input metrics	AICc model input metrics
<i>No. saplings per 30x30m</i>	-	(1) ht_p35_2 (points-all) [leaf-on] (2) ht_p50 (points-all) [leaf-off] (3) wd_p35_2 (width-all) [leaf-on]
<i>No. native saplings per 30x30m</i>	-	(1) ht_p35_2 (points-all) [leaf-on] (2) medht (points-all) [leaf-off] (3) wd_p35_2 (width-all) [leaf-on]
<i>Percentage of native saplings per 30x30m</i>	(1) skwht (points-all) [leaf-off] (2) Domht (points-all) [leaf-off]	(1) meanamp (amplitude-all) [leaf-off] (2) wd_p40 (points-all) [leaf-off]
<i>No. of sapling species per 30x30m</i>	(1) ht_p35_2 (points-all) [leaf-on] (2) varamp (amplitude-all) [leaf-off]	(1) ht_p40_2 (points-all) [leaf-on] (2) TIF_TreeNo (ITC) (3) varamp (amplitude-all) [leaf-off] (4) wd_p25 (width-all) [leaf-off]
<i>No. seedlings per 30x30m</i>	(1) adevamp_2 (amplitude-all) [leaf-on]	(1) medgamp_2 (amplitude-ground) [leaf-on] (2) varngamp_2 (amplitude-ground) [leaf-on]
<i>No. native seedlings per 30x30m</i>	(1) skwngamp (amplitude-non-ground) [leaf-off] (2) vargwid (width-ground) [leaf-off] (3) skwgwid_2 (width-ground) [leaf-on]	(1) skwngamp (amplitude-non-ground) [leaf-off] (2) vargwid (width-ground) [leaf-on]
<i>Percentage of native seedlings per 30x30m</i>	(1) medgamp_2 (amplitude-ground) [leaf-on] (2) ht_p25_2 (points-all) [leaf-on]	(1) skwngamp (amplitude-non-ground) [leaf-off] (2) vargwid (width-ground) [leaf-off]
<i>No. of seedling species per 30x30m</i>	(1) skwgamp_2 (amplitude-ground) [leaf-on]	(1) medgamp (amplitude-ground) [leaf-off] (2) stdgwid (width-ground) [leaf-off]
<i>Seedlings Shannon index for native species</i>	-	-
<i>No. vascular species</i>	(1) varnght_2 (points-non-ground) [leaf-on]	(1) varnght_2 (points-non-ground) [leaf-on]
<i>No. bryophyte species</i>	(1) kurngamp (amplitude-non-ground) [leaf-off] (2) lon_can_layer (points-all) [leaf-on] (3) TIF_STD (ITC)	(1) skwngamp (amplitude-non-ground) [leaf-off]
<i>Percentage of bare soil within 30x30m</i>	(1) ht_p20_2 (points-all) [leaf-on] (2) loff_can_depth (points-all) [leaf-off] (3) TIF_STD (ITC)	(1) ht_p15_2 (points-all) [leaf-on] (2) loff_can_depth (points-all) [leaf-off] (3) meannght_2 (points-non-ground) [leaf-on]

7.4.5 Metrics required for condition index construction

The two metrics required for inputs into the condition indices and their associated R^2 values are listed in Table 7.26. The sum of the DBH differences has a higher R^2 value for OLS, whereas the opposite is true for estimates of the models for the index of vertical separation. The AICc evidence ratio indicates only small increases in the first ‘best’ model over the second.

FW LiDAR regression model inputs are given in Table 7.27. The FW inputs for these models vary between OLS and AICc. The FW metrics for the sum of the DBH differences for OLS is based primarily on height-points, whereas curve-width is used more in AICc. The models for the index of vertical separation both begin with a point-height metric relating to the distribution of values, but then differ, using amplitude or width metrics. Both leaf-on and leaf-off metrics were used.

Table 7.26 – Condition index construction metrics, R^2 and AICc weights (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	OLS R^2	AICc R^2	AICc weight	AICc of 2 nd best	AICc evidence ratio
<i>Sum of the DBH differences between neighbouring trees (cm)</i>	<u>0.912</u>	0.714	0.264	0.244	1.082
<i>Index of vertical separation</i>	0.467	<u>0.601</u>	0.074	0.056	1.321

Table 7.27 – FW LiDAR metric inputs for each forest condition index metric regression model (Appendix C.3, C.6-C.8 should be consulted for a description of FW LiDAR metric names)

Metric Name:	OLS model input metrics	AICc model input metrics
<i>Sum of the DBH differences between neighbouring trees (cm)</i>	(1) TIF_TreeNo (ITC) (2) kurht (points-all) [leaf-off] (3) kurnght_2 (points-non-ground) [leaf-on] (4) kurngwid_2 (width-non-ground) [leaf-on]	(1) adevwid (width-all) [leaf-off] (2) amp_p90 (amplitude-all) [leaf-off] (3) ht_p20_2 (points-all) [leaf-on] (4) kurngwid_2 (width-non-ground) [leaf-on]
<i>Index of vertical separation</i>	(1) skwnght_2 (points-non-ground) [leaf-on] (2) amp_p85 (amplitude-all) [leaf-off]	(1) varht_2 (points-all) [leaf-on] (2) wd_p30 (width-all) [leaf-off]

7.5 Validation of FW LiDAR remote sensing derived estimates of forest variables against field data

This results section outlines the various direct and statistical model estimates in comparison with validation fieldwork measurements. It begins with the direct comparisons which can be made using FW LIDAR remote sensing data. This is followed by a summary of the various statistical model estimates derived from remote sensing means, how they are calculated and their relationships to the validation field measurements. RMSE and NRMSE values are presented to illustrate the accuracy of each model prediction. NRMSE is a modified RMSE value to account for the field data population range, and produces a value of 0 to 1. NRMSE values of 0.5 or greater indicate poor model performance.

7.5.1 Direct comparisons between field measurements and remote sensing metrics

A number of direct comparisons can be made between field measures and remote sensing measures; these are: (i) average tree height per plot; (ii) number of canopy layers; (iii) and canopy openness. It should be noted that TIFFS ITC metrics produced for the FW CHM were poorer to those produced for the DR CHM due to the reduced pulse rate and lower horizontal sample resolution and were removed.

Area-based estimates of average tree height per plot and the number of canopy layers is presented in Table 7.28. Dominant height and the number of canopy layers were produced using the SPDlib software for both leaf-on and leaf-off FW data. Dominant height for leaf-on data had the best correspondence with Lorey's mean height field data, producing the lowest RMSE and NRMSE values of 3.65m and 0.19 respectively. The number of canopy layers produced similar RMSE and NRMSE values for both leaf-on and leaf-off datasets, but with the leaf-off RMSE being slightly lower.

Estimates for canopy openness were calculated using the SPDlib software for FW leaf-on and leaf-off datasets. A comparison between FW and field estimates is presented in Table 7.29. The lowest RMSE value, at 8.28%, was calculated for the FW leaf-on estimates of canopy openness.

Table 7.28 – Comparison of forest structural metrics from FW LiDAR data with field data

Site no.	Lorey's Height (m) [Field]	[FW] Dom. Height (m) [leaf-off]	[FW] Dom. Height (m) [leaf-on]	No. Canopy Layers [Field]	[FW] no. of canopy layers [leaf-off]	[FW] no. of canopy layers [leaf-on]
22	12.83	18.05	18.33	2	2	3
23	16.85	11.95	12.25	2	2	2
24	19.17	21.14	21.38	2	2	3
25	22.13	22.89	23.34	2	3	3
26	21.05	17.44	17.54	1	3	3
27	24.79	29.73	29.44	3	4	5
28	23.35	18.76	19.09	1	3	3
29	17.89	19.34	22.03	2	3	4
30	19.38	19.77	20.74	2	2	3
31	28.78	30.31	31.31	1	2	2
32	25.64	21.72	21.80	1	3	3
33	28.36	11.13	26.76	1	2	2
34	22.99	23.89	24.02	3	3	3
35	23.89	27.22	27.84	3	3	3
36	25.12	22.42	22.17	2	3	2
37	17.91	21.79	22.03	2	2	3
38	21.4	17.86	18.20	3	2	3
39	33.12	26.74	26.98	3	3	3
40	24.45	22.07	22.31	3	3	3
41	18.99	22.56	23.36	2	4	3
	RMSE	5.182	3.652	RMSE	1.072	1.180
	NRMSE	0.273	0.192	NRMSE	0.357	0.390

Table 7.29 – Canopy Openness comparison

Site no.	Canopy Openness (%) [Field]	Canopy openness (%) (Leaf off) [FW]	Canopy openness (%) (Leaf on) [FW]
22	20.59	32.52	25.29
23	17.24	32.13	26.66
24	19.59	36.68	30.53
25	21.01	33.94	24.30
26	22.81	37.76	32.94
27	12.97	27.93	18.35
28	26.84	40.40	36.71
29	22.76	33.62	27.98
30	15.68	29.25	17.37
31	14.67	29.62	15.59
32	15.8	32.11	19.71
33	18.89	45.94	31.24
34	15.72	34.74	26.28
35	21.33	39.13	36.05
36	14.75	29.31	10.80
37	23.87	37.57	30.03
38	17.72	33.29	30.36
39	22.8	42.51	33.61
40	16.21	25.67	13.98
41	10.35	35.55	16.66
	RMSE	16.45	8.28

7.5.2 Validation of remote sensing derived statistical models against field data

In this section the results of the regression models derived from FW LiDAR datasets were compared with the 2012 field data measurements. RMSEs and NRMSEs were calculated for each of the remote sensing derived models; see Table 7.30 for overstorey results. As before, statistical methods performed better than the direct measurements provided by TIFFS for the number of trees and the NN spacing of stems. The direct measurements of the average crown horizontal area, and total crown horizontal area performed as well as or better than statistical estimates. NRMSE values for three field metrics for both OLS and AICc models were above 0.5; these were: mean DBH, percentage very big trees, and SH index (by basal area). The NRMSE value for the AICc model for the standard deviation of tree NN spacing was above 1 (i.e. greater than 100% of the range of the field values). Of the remaining 17 metrics, 10 AICc RMSE values were lower than those of their OLS counterparts.

The values reported in Table 7.31 refer to understorey structure and composition. The models relating to percentage native saplings and the number of bryophyte species had NRMSE values above 0.5 for both OSL and AICc. The AICc model for estimating the number of saplings produced an NRMSE value above 6 (i.e. far above the field value range). Of the 14 remaining models, AICc models had lower RMSE values than OLS models for 9 of the field metrics.

Table 7.30 – The OLS and AICc FW LiDAR regression models. Part 1: overstorey composition.

Field metric	ITC RMSE*	ITC NRMSE*	OLS RMSE	OLS NRMSE	AICc RMSE	AICc NRMSE
<i>Number of trees per 30x30m</i>	24.077	0.325	15.974	0.216	20.071	0.271
<i>Number of native trees per 30x30m</i>	-	-	13.927	0.268	13.927	0.268
<i>Percentage of native trees</i>	-	-	35.058	0.351	29.741	0.297
<i>No. of tree species per 30x30m</i>	-	-	2.145	0.357	1.658	0.276
<i>Average tree NN spacing (m)</i>	4.234	1.275	1.321	0.398	2.059	0.620
<i>STDev. of tree NN spacing (m)</i>	0.657	0.790	0.394	0.474	0.953	1.147
<i>Mean DBH (cm)</i>	-	-	17.125	0.588	15.406	0.529
<i>STDev. of DBH (cm)</i>	-	-	11.455	0.235	9.100	0.187
<i>Basal area within 30x30m (m²)</i>	-	-	0.906	0.283	1.155	0.360
<i>(i) Percentage big trees (DBH 40>x<80cm)</i>	-	-	26.629	0.380	31.256	0.447
<i>(ii) Percentage very big trees (DBH >80cm)</i>	-	-	10.306	1.701	17.322	2.858
<i>(iii) Percentage big trees (DBH >40cm)</i>	-	-	21.888	0.313	25.667	0.367
<i>Shannon index (by stem count)</i>	-	-	0.577	0.403	0.623	0.435
<i>Simpson index (by stem count)</i>	-	-	0.289	0.224	0.285	0.221
<i>Shannon index (by BA)</i>	-	-	0.604	0.786	0.499	0.649
<i>Simpson index (by BA)</i>	-	-	0.846	0.293	0.795	0.275
<i>Average height of the first live branch (m)</i>	-	-	2.592	0.239	2.657	0.245
<i>STDev. of the height of the first live branch (m)</i>	-	-	2.221	0.431	2.317	0.449
<i>Average crown horizontal area (m²)</i>	25.459	0.292	24.387	0.280	34.520	0.396
<i>STDev. crown horizontal area (m²)</i>	-	-	27.897	0.488	17.594	0.308
<i>Total crown horizontal area (m²)</i>	558.118	0.387	645.640	0.447	671.700	0.465

*Denotes this was not derived statistically and cannot be applied to all metrics

Table 7.31 – The OLS and AICc FW LiDAR regression models. Part 2: understorey composition.

Field metric	OLS RMSE	OLS NRMSE	AICc RMSE	AICc NRMSE
<i>Standing dead wood volume (m³)</i>	1.360	0.165	<u>1.360</u>	<u>0.165</u>
<i>Standing dead wood decay class</i>	0.203	0.304	<u>0.193</u>	<u>0.290</u>
<i>Downed dead wood volume (m³)</i>	<u>2.490</u>	<u>0.274</u>	2.903	0.319
<i>Downed dead wood decay class</i>	0.221	0.471	<u>0.203</u>	<u>0.432</u>
<i>No. saplings per 30x30m</i>	-	-	26.461	0.245
<i>No. native saplings per 30x30m</i>	-	-	11.465	0.302
<i>Percentage of native saplings per 30x30m</i>	75.162	0.752	63.774	0.638
<i>No. of sapling species per 30x30m</i>	<u>1.424</u>	<u>0.475</u>	20.877	6.959
<i>No. seedlings per 30x30m</i>	<u>331.964</u>	<u>0.355</u>	337.426	0.360
<i>No. native seedlings per 30x30m</i>	<u>245.151</u>	<u>0.340</u>	269.858	0.375
<i>Percentage of native seedlings per 30x30m</i>	<u>39.719</u>	<u>0.397</u>	43.260	0.433
<i>No. of seedling species per 30x30m</i>	3.534	0.272	<u>3.098</u>	<u>0.238</u>
<i>Seedlings Shannon index for native species</i>	-	-	-	-
<i>No. vascular species</i>	<u>1.698</u>	<u>0.283</u>	1.844	0.307
<i>No. bryophyte species</i>	9.161	0.916	8.365	0.837
<i>Percentage of bare soil within 30x30m</i>	<u>39.822</u>	<u>0.419</u>	39.955	0.421
<i>Sum of the DBH differences between neighbouring trees (cm)</i>	<u>0.130</u>	<u>0.257</u>	0.287	0.568
<i>Index of vertical separation</i>	5.486	0.184	<u>2.762</u>	<u>0.092</u>

7.6 Estimating or calculating metrics required for input into the complex stand diversity index (CSDI)

A number of additional metrics were required for the construction of the complex stand diversity index (CSDI). The CSDI is composed of four component indices relating to different forest stand attributes. The first of these indicates the species diversity within a specific plot (see chapter 6.3.1). The second index is an indication of how much tree diameters (i.e. DBH) differs within the plot, which was estimated through statistical regression. The third of these component indices relates to the distribution of tree stems within a plot, for the three smallest and three largest distances between NN trees in metres. The TIFFS software and a custom R script was utilised to extract these distance metrics from ITC centroid locations using LiDAR data. Table 7.32 summarises the three minimum LiDAR derived TIFFS ITC and corrected LiDAR TIFFS ITC derived NN distances against the field measured NN distances. Table 7.33 summarises the three maximum NN distances. Comparing field NN distances with TIFFS ITC centroid distances produced estimates with RMSE values of approximately $\pm 4\text{m}$ for minimum distances and between ± 4 and $\pm 6\text{m}$ for maximum distances. RMSE values for corrected TIFFS ITC centroid distances are significantly lower ranging from $\pm 1.66\text{-}1.90\text{m}$ for minimum NN distances, and $\pm 2.15\text{-}3.35\text{m}$ for maximum distances. Note that corrected LiDAR TIFFS ITC derived NN distances were calculated by subtracting twice the standard deviation of tree NN spacing from the TIFFS distances. The value of the standard deviation of tree NN spacing was derived statistically, the model producing the smallest RMSE/NRMSE overall was chosen to provide this estimate, see Chapter 8.4.

The fourth component of the CSDI index requires a calculation using the two smallest and two largest tree crown horizontal diameters, height to the first live branch (or crown base) and the number of trees within the plot. The first input, relating to crown diameters, could be approximated using data from DR LiDAR derived ITC objects. A comparison of field measured crown diameters and ITC object attributes is summarised in Table 7.34. RMSEs for minimum crown diameters range from $\pm 1.77\text{-}2.07\text{m}$ and for maximum crown diameters from $\pm 2.07\text{-}4.01\text{m}$. The second input required for the fourth component index was an estimate of the lowest crown base height (i.e. the smallest distance from the ground to the first live branch). This value was approximated through the use of the statistically derived estimates of the average height of the first live branch and subtracting the standard deviation of the height of the first live branch (see Chapter 8.4). This calculation was applied to both DR and FW LiDAR, and the results are presented in Table 7.35. The lowest RMSE value (2.40m) is

recorded for the models produced by combining DR LiDAR derived models. The FW derived models produced the highest RMSE value (3.50m).

Table 7.32 – A summary of three minimum recorded NN distances for field measurements, LiDAR derived TIFFS ITC and corrected TIFFS ITC measurements.

Site no.	Min. NN distance #1 (m) [Field]	Min. NN distance #1 (m) [TIFFS ITC]	Min. NN distance #1 (m) [Corr. ITC]	Min. NN distance #2 (m) [Field]	Min. NN distance #2 (m) [TIFFS ITC]	Min. NN distance #2 (m) [Corr. ITC]	Min. NN distance #3 (m) [Field]	Min. NN distance #3 (m) [TIFFS ITC]	Min. NN distance #3 (m) [Corr. ITC]
23	0.60	4.45	1.47	0.60	4.45	1.47	0.73	4.54	1.56
24	0.50	4.91	2.81	0.50	4.91	2.80	0.53	5.43	3.32
25	0.15	5.45	2.72	0.15	5.45	2.72	0.30	5.56	2.84
26	0.40	6.86	2.71	0.40	6.86	2.71	0.85	7.10	2.95
27	0.32	6.82	4.25	0.49	6.82	4.25	0.59	7.10	4.53
28	1.28	4.40	0.84	2.08	4.40	0.84	2.20	6.32	2.76
29	0.22	5.00	1.96	0.40	5.00	1.96	0.40	6.93	3.89
30	0.52	4.03	2.14	0.52	4.03	2.14	0.58	4.48	2.58
31	2.59	3.16	2.19	2.59	3.16	2.19	2.98	3.22	2.24
32	2.12	4.69	1.36	2.67	4.69	1.36	2.87	5.81	2.49
33	2.00	3.99	2.90	2.50	3.99	2.90	2.60	4.26	3.18
34	0.95	5.75	2.95	0.95	5.75	2.95	1.40	5.89	3.10
35	1.69	4.39	2.79	1.69	4.39	2.79	1.83	5.42	3.81
36	2.38	6.54	3.92	2.53	6.54	3.92	2.54	7.12	4.51
37	1.87	5.89	2.91	1.87	5.89	2.91	2.12	6.79	3.82
38	2.05	4.52	1.84	2.05	4.52	1.84	2.06	4.84	2.16
39	3.42	5.31	2.45	3.42	5.31	2.45	3.66	6.15	3.29
40	1.40	5.63	2.82	1.48	5.63	2.82	3.45	6.12	3.31
41	0.97	5.08	1.73	0.97	5.08	1.73	1.11	6.05	2.70
-	RMSE	4.04	1.67	RMSE	3.95	1.66	RMSE	4.32	1.90

Table 7.33 – A summary of three maximum recorded NN measurements for field measurements, LiDAR derived TIFFS ITC and corrected TIFFS ITC measurements.

Site no.	Max. NN distance #1 (m) [Field]	Max. NN distance #1 (m) [TIFFS ITC]	Max. NN distance #1 (m) [Corr. ITC]	Max. NN distance #2 (m) [Field]	Max. NN distance #2 (m) [TIFFS ITC]	Max. NN distance #2 (m) [Corr. ITC]	Max. NN distance #3 (m) [Field]	Max. NN distance #3 (m) [TIFFS ITC]	Max. NN distance #3 (m) [Corr. ITC]
23	4.06	7.33	4.35	4.43	7.32	4.34	8.28	7.32	4.34
24	5.28	10.59	8.48	6.80	10.32	8.21	7.02	10.32	8.21
25	3.95	12.36	9.64	5.20	9.67	6.95	5.60	9.67	6.95
26	4.18	11.62	7.47	4.20	10.18	6.02	4.60	10.18	6.02
27	3.34	12.08	9.50	3.54	11.68	9.11	3.57	11.68	9.11
28	4.54	11.93	8.36	4.54	11.93	8.36	5.44	11.93	8.36
29	4.17	11.80	8.76	4.74	11.05	8.02	4.95	11.05	8.02
30	3.20	8.78	6.89	3.21	8.59	6.69	4.20	8.59	6.69
31	5.79	6.98	6.00	6.45	6.80	5.82	7.84	6.80	5.82
32	5.63	12.85	9.52	6.62	9.74	6.41	7.27	9.74	6.41
33	5.60	6.60	5.52	5.90	6.55	5.47	5.90	6.55	5.47
34	5.81	12.19	9.40	5.85	9.98	7.18	5.98	9.98	7.18
35	5.86	10.45	8.84	7.00	7.92	6.32	7.35	7.92	6.32
36	5.74	8.91	6.30	5.80	8.32	5.71	6.06	8.32	5.71
37	5.51	10.99	8.02	6.45	10.84	7.87	6.50	10.84	7.87
38	4.60	11.23	8.55	5.07	8.98	6.29	6.75	8.98	6.29
39	6.10	11.49	8.63	8.05	11.02	8.16	9.20	11.02	8.16
40	6.08	10.95	8.14	6.10	10.95	8.14	6.25	10.95	8.14
41	4.74	9.22	5.87	4.75	8.95	5.60	5.14	8.95	5.60
-	RMSE	5.89	3.34	RMSE	4.50	2.15	RMSE	4.08	2.18

Table 7.34 – A summary of the two minimum and two maximum tree crown horizontal diameters for field measurement and TIFFS ITC measurements.

Site no.	Min. Crown diameter #1 (m) [Field]	Min. Crown diameter #1 (m) [TIFFS ITC]	Min. Crown diameter #2 (m) [Field]	Min. Crown diameter #2 (m) [TIFFS ITC]	Max. Crown diameter #1 (m) [Field]	Max. Crown diameter #1 (m) [TIFFS ITC]	Max. Crown diameter #2 (m) [Field]	Max. Crown diameter #2 (m) [TIFFS ITC]
22	6.21	3.39	6.40	3.74	17.45	17.81	21.35	16.43
23	1.03	3.19	1.18	3.57	6.48	9.51	7.19	8.44
24	1.40	3.74	1.51	3.74	9.43	14.71	11.65	12.51
25	5.13	2.76	5.29	4.22	16.05	16.62	18.55	14.63
26	2.99	4.22	3.94	6.58	11.77	16.27	13.18	14.00
27	2.49	2.99	2.94	3.74	11.57	24.33	13.59	13.11
28	3.47	4.92	3.88	5.17	9.65	13.87	9.86	13.40
29	2.69	4.65	2.77	7.05	12.31	17.08	16.02	12.82
31	3.50	1.95	5.18	1.95	9.62	9.64	9.66	8.06
32	4.99	4.92	6.22	4.92	13.23	16.23	14.74	11.83
33	1.00	2.52	2.73	3.57	7.56	8.59	8.18	8.52
34	3.38	5.41	4.54	5.41	10.21	11.17	13.50	11.06
35	3.11	3.91	3.53	4.65	10.53	11.23	12.68	10.52
36	4.58	5.29	4.69	5.64	8.13	10.09	8.26	9.84
37	4.69	2.76	4.96	6.08	12.60	14.41	13.69	12.05
38	4.01	3.19	4.03	4.07	12.23	16.04	12.36	11.17
39	5.07	2.26	5.30	4.65	10.20	13.21	13.35	12.36
40	5.81	4.22	7.55	4.22	12.98	16.55	20.12	12.87
41	3.65	3.39	4.20	3.74	13.78	13.26	13.80	12.67
-	RMSE	1.77	RMSE	2.07	RMSE	2.77	RMSE	4.01

Table 7.35 – A summary of approximated lowest crown base height value

Site no.	Min. height of crown base (m) [Field]	Min. height of crown base (m) [DR]	Min. height of crown base (m) [FW]
22	1.25	5.26	4.87
23	2.00	2.78	3.99
24	1.75	3.53	1.69
25	2.00	4.40	3.75
26	2.00	2.85	0.48
27	0.75	0.83	0.36
28	5.05	3.44	4.47
29	1.50	2.37	2.61
30	-	-	-
31	3.98	4.20	4.74
32	2.00	3.35	3.76
33	1.50	8.76	13.06
34	1.00	1.94	4.98
35	3.20	4.83	4.91
36	2.50	5.12	6.93
37	2.00	4.44	5.08
38	1.70	1.95	2.19
39	4.40	3.80	5.84
40	2.20	4.38	5.14
41	1.00	2.91	4.51
-	RMSE	2.40	3.50

7.7 Summary of findings

The comparisons of elevations from DTMs produced from DR and FW LiDAR with field-based GPS and total station measurements produced very similar results overall. However, slightly better accuracy was achieved using leaf-off and FW data.

In a comparison of DR and FW point cloud datasets, FW provides a significantly higher proportion of points through the fitting of Gaussian peaks to each of the laser waveforms. A 143% increase in points relative to initial pulses was found for FW, whereas for DR this was only 30%, which indicates that the majority of the DR pulses only supplied a single return. FW data produced a greater number of pulses at all levels within the mid-canopy relative to DR data. Due to the nature of the sampling of the LiDAR datasets DR data provided a higher density and more evenly distributed 3D points whereas FW data provided more points vertically per pulse, but were subject to greater spacing horizontally through all vertical levels.

For DR LiDAR automated ITC detection underestimated the total tree stem number against counts in validation field plots; however they did more closely resemble over-storey tree stem counts. Average tree crown height resembled field Lorey's average height to within

$\pm 3.46\text{m}$ (RMSE). The total number of canopy layers derived from LiDAR data could be up to one layer away from the field measured value. Estimates for canopy openness were best for leaf-on data, and could be out by up to 13.3% (RMSE).

DR derived regression models could be produced for all thirty-nine indicator metrics. The OLS regression approach could not produce a model for predicting two indices of species diversity, namely the SI (by stem count) and the SH (by basal area). RMSE and NRMSE values were computed for the OLS and AICc models for predicting field-level metrics. When comparing model predictions against field collected data six of the thirty-nine models produced NRSME above 0.5. A total of 30 metrics had NRMSE values below 0.4, and 19 of the metrics had NRMSE values below 0.3. Several of the metrics with higher NRMSE (>0.4) were related to understorey composition, including the number of tree sapling species. Average crown area and total crown area were better predicted from ITC information. For DR OLS and AICc there is no definitive better method.

For FW LiDAR automated ITC detection was identical to that produced through DR LiDAR. Field-plot derived Lorey's average height corresponded best with dominant tree crown height when using leaf-on data and was within $\pm 3.65\text{m}$ (RMSE). Again, the number of canopy layers could be within ± 1 of the field-measured value. Canopy openness of field data corresponded best with leaf-on FW LiDAR, within $\pm 8\%$ (RMSE).

FW derived regression models could be completed for all but one of the thirty-nine indicator metrics. The metric Seedlings Shannon Index for native species could not be modelled. OLS regression approaches could not model the number of saplings or the number of native saplings. RMSE and NRMSE values were computed for the OLS and AICc regression models for predicting field-level metrics. When comparing model predictions against field collected data, six of the thirty-nine models produced NRSME above 0.5. A total of 27 metrics had NRMSE values below 0.4, and 18 of the metrics had NRMSE values below 0.3. Again, many higher NRMSE values (>0.4) were found when considering understorey composition. Estimates of average crown area and total crown area from ITC data were better than statistically derived metrics. As with DR, for FW OLS and AICc there is no definitive optimal method.

When considering the entirety of the models produced by DR and FW LiDAR metrics a total of four field metrics could not be predicted, where the NRMSE value was higher than 0.5. Generally, NRMSE values were very close between model estimates from DR and FW models. DR predicative models produced the lowest NRMSE values for 17 of the 35 metrics, whereas FW models produced the lowest NRMSE values for 18 of the 35 metrics. This pattern is repeated when the best model (with NRMSE <0.3) is removed, i.e. out of a total of 22 field-level metrics 10 are best estimated by DR and 12 are best estimated by FW.

A total of 27 metrics could be predicted with an R^2 value above 0.5, and NRMSE value lower than 0.35. In 14 of these models DR LiDAR produced lower NRMSE values, whereas for FW this was 13.

Two component inputs into the CSDI, condition assessments were extracted from the TIFFS ITC dataset in addition to other statistically derived attributes. The relevant ITC metrics involving the nearest neighbour (NN) distances of tree stems produced estimates with RMSE for the smallest distances within ± 1.66 -1.90m and within ± 2.15 -3.35m for the largest distances when a correction was applied. Estimates of tree crown diameter produced RMSE values for minimum diameters of ± 1.77 -2.07m and ± 2.07 -4.01m for the largest diameters.

Estimates of the lowest crown base height were made by subtracting the standard deviation of the height of the first live branch from the average height to the first live branch using metrics statistically derived from DR and FW metrics. DR LiDAR produced the estimates with the smallest RMSE value (2.40m) and thus was used in the CSDI calculation.

Chapter 8 – Forest assessment using a fusion between hyperspectral and LiDAR datasets

The combination of hyperspectral and LiDAR remote sensing metrics to estimate field plot-level information is explored in this chapter. The chapter begins with an overview of the statistical models generated by combining hyperspectral and LiDAR data, and each model's validation against field data. Section 8.1 describes the models derived by combining hyperspectral data with DR LiDAR, and then with FW LiDAR. Section 8.2 compares the performance of the models with the field validation data from 2012. Section 8.3 reviews and summaries all the regression models derived for the remote sensing datasets. Section 8.4 reviews and summaries the predictive capacity of the models and identifies the “best” models for each field metric. Section 8.5 provides an overall summary of the chapter.

8.1 Estimating field-level forest structural and compositional metrics from combined hyperspectral and LiDAR data

The two regression approaches (OLS and AICc) were applied to the two composite hyperspectral and LiDAR datasets, for each of the 39 field metrics. Each of the models presented was statistically significant at $p < 0.05$. The inputs can be summarised as those statistics derived from ITC objects and area-based statistical summaries of hyperspectral and LiDAR data layers. What follows is a breakdown of the regression models produced by forest structural, compositional, deadwood, and regeneration and understorey metrics, and finally the two additional metrics required for index construction which are described later (Chapter 9).

8.1.1 Forest structure within the plot

A listing of the 14 forest structural metrics and their associated R^2 values is given in Table 8.1. The first of the combinations, hyperspectral and DR LiDAR, produced OLS and AICc models with high R^2 values. The majority of these values are above 0.7. For only three of these models did OLS methods produce higher R^2 values in comparison with AICc. The second combination concerned hyperspectral and FW LiDAR, which again produced a high proportion of R^2 values above 0.7. In this case, OLS produced five models with a higher R^2 value than the AICc counterpart.

In comparing the R^2 values between the two sets of composite models, eight of the total fourteen models for hyperspectral and DR had higher R^2 values than hyperspectral combined with FW metrics, but the values were very similar.

The inputs to the regression models for the combination of hyperspectral and DR LiDAR are listed in Table 8.2. Table 8.3 documents the inputs for the regression models for a combination of hyperspectral and FW LiDAR. Some models use LiDAR or hyperspectral data, rather than a combination of both. For the hyperspectral and DR composite dataset, 7 of 14 models use only LiDAR metrics for OLS and 4 for AIC, while only 1 uses hyperspectral only for OLS and 2 for AIC. For the hyperspectral and FW composite, 7 of 14 models use only LiDAR metrics for OLS, and 5 for AIC, whereas hyperspectral is always used in a combination. The majority of models also utilise both leaf-on and leaf-off data.

Table 8.1 – Forest structural metrics R^2 for hyperspectral and LiDAR combinations (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	DR and Hyper. OLS R^2	DR and Hyper. AICc R^2	FW and Hyper. OLS R^2	FW and Hyper. AICc R^2
Number of trees per 30x30m	0.804	<u>0.912</u>	0.699	<u>0.771</u>
Average tree NN spacing (m)	0.785	<u>0.854</u>	0.785	<u>0.819</u>
STDev. Of tree NN spacing (m)	0.771	<u>0.771</u>	<u>0.835</u>	0.664
Mean DBH (cm)	0.664	<u>0.755</u>	0.796	<u>0.796</u>
STDev. of DBH (cm)	0.779	<u>0.867</u>	0.692	<u>0.819</u>
Basal area within 30x30m	<u>0.761</u>	0.537	0.573	<u>0.822</u>
Percentage big trees (DBH 40>x<80cm)	0.850	<u>0.925</u>	0.809	<u>0.915</u>
Percentage very big trees (DBH >80cm)	0.654	<u>0.799</u>	0.592	<u>0.744</u>
Percentage big trees (DBH >40cm)	0.924	<u>0.942</u>	<u>0.915</u>	0.908
Average height of the first live branch (m)	0.820	<u>0.876</u>	<u>0.876</u>	0.805
STDev. of the height of the first live branch (m)	<u>0.808</u>	0.629	0.743	<u>0.805</u>
Average crown horizontal area (m^2)	0.852	<u>0.859</u>	<u>0.860</u>	0.828
STDev. Crown horizontal area (m^2)	<u>0.818</u>	0.727	<u>0.892</u>	0.796
Total Crown horizontal area (m^2)	0.601	<u>0.662</u>	0.605	<u>0.761</u>

Table 8.2 – Hyperspectral and DR LiDAR metric inputs for each forest structural metric regression model (Appendix B and C should be consulted for a description of remote sensing metric names.)

Metric Name:	DR and Hyper. OLSmodel input metrics		DR and Hyper. AICc model input metrics	
<i>Number of trees per 30x30m</i>	TIF_m_area p020_2	(LiDAR) [ITC] (LiDAR) [leaf-on]	cri1_Sum SHA_Kur_2 SHA_Var_2 TIF_m_area	(Hyper.) [leaf-off] (LiDAR) [leaf-on] (LiDAR) [leaf-on] (LiDAR)
<i>Average tree spacing (m)</i>	TIF_Space sri2_STD ari2_Sum	(LiDAR) [ITC] (Hyper.) [leaf-on] (Hyper.) [leaf-on]	rgri2_Range SHA_MAD_2 TIF_Space TIF_to_area	(Hyper.) [leaf-on] (LiDAR) [leaf-on] (LiDAR) [ITC] (LiDAR) [ITC]
<i>STDev. Of tree spacing (m)</i>	TIF_m_area p060	(LiDAR) [ITC] (LiDAR) [leaf-off]	no_tree pri2_STD sri2_Range	(Hyper.) [ITC] (Hyper.) [leaf-on] (Hyper.) [leaf-on]
<i>Mean DBH (cm)</i>	TIF_m_area	(LiDAR) [ITC]	rendvi1_STD TIF_Space	(Hyper.) [leaf-off] (LiDAR) [ITC]
<i>STDev. of DBH (cm)</i>	mresri1_STD SHN_Ske	(Hyper.) [leaf-off] (LiDAR) [leaf-off]	rendvi1_STD SHN_Ske TIF_TreeNo	(Hyper.) [leaf-off] (LiDAR) [leaf-off] (LiDAR) [leaf-off]
<i>Basal area within 30x30m</i>	SI mresri1_STD sri1_STD TIF_Space	(Hyper.) [ITC] (Hyper.) [leaf-off] (Hyper.) [leaf-off] (LiDAR) [ITC]	arvi1_Range SI	(Hyper.) [leaf-off] (Hyper.) [ITC]
<i>Percentage big trees (DBH 40>x<80cm)</i>	SHA_MAD_2 TIF_Space	(LiDAR) [leaf-on] (LiDAR) [ITC]	rgri2_Max SHA_Var_2 TIF_Space	(Hyper.) [leaf-on] (LiDAR) [leaf-on] (LiDAR) [ITC]
<i>Percentage very big trees (DBH >80cm)</i>	vrei1_STD	(Hyper.) [leaf-off]	p095i no_tree sri2_STD	(LiDAR) [leaf-off] (Hyper.) [ITC] (Hyper.) [leaf-on]
<i>Percentage big trees (DBH >40cm)</i>	SHA_STD_2 TIF_Space	(LiDAR) [leaf-on] (LiDAR) [ITC]	SHA_STD_2 TIF_Space TIF_TreeNo	(LiDAR) [leaf-on] (LiDAR) [ITC] (LiDAR) [ITC]
<i>Average height of the first live branch (m)</i>	loff_can_depth SHN_Mean	(LiDAR) [leaf-off] (LiDAR) [leaf-off]	loff_max_gap TIF_Space	(LiDAR) [leaf-off] (LiDAR) [ITC]
<i>STDev. of the height of the first live branch (m)</i>	TIF_m_area TIF_meanCV sri2_Max SI	(LiDAR) [ITC] (LiDAR) [ITC] (LiDAR) [leaf-on] (Hyper.) [ITC]	lon_max_gap TIF_to_area	(LiDAR) [leaf-on] (LiDAR) [ITC]
<i>Average crown horizontal area (m²)</i>	TIF_m_area mresri2_Max TIF_meanCV	(LiDAR) [ITC] (Hyper.) [leaf-on] (LiDAR) [ITC]	mresri2_Max SIN_Kur TIF_m_area	(Hyper.) [leaf-on] (LiDAR) [leaf-off] (LiDAR) [ITC]
<i>STDev. Crown horizontal area (m²)</i>	TIF_m_area SIA_Med	(LiDAR) [ITC] (LiDAR) [leaf-off]	SIN_Med sri2_STD TIF_m_area	(LiDAR) [leaf-off] (Hyper.) [leaf-on] (LiDAR) [ITC]
<i>Total Crown horizontal area (m²)</i>	ari2_Sum SIA_Var_2	(Hyper.) [leaf-on] (LiDAR) [leaf-on]	SHA_Mean_2 SHA_STD_2 SIA_Var_2	(LiDAR) [leaf-on] (LiDAR) [leaf-on] (LiDAR) [leaf-on]

Table 8.3 – Hyperspectral and FW LiDAR metric inputs for each forest structural metric regression model (Appendix B and C should be consulted for a description of remote sensing metric names.)

Metric Name:	FW and Hyper. OLSmodel input metrics		FW and Hyper. AICc model input metrics	
<i>Number of trees per 30x30m</i>	TIF_m_area ht_p20_2	(LiDAR) [ITC] (LiDAR) [leaf-on]	kurht_2 sri2_Min TIF_m_area	(LiDAR) [leaf-on] (Hyper.) [leaf-on] (LiDAR) [ITC]
<i>Average tree spacing (m)</i>	TIF_Space sri2_STD ari2_Sum	(LiDAR) [ITC] (Hyper.) [leaf-on] (Hyper.) [leaf-on]	pri2_STD TIF_Space varht_2	(Hyper.) [leaf-on] (LiDAR) [ITC] (LiDAR) [leaf-on]
<i>STDev. of tree spacing (m)</i>	TIF_m_area ht_p40 kurngwid	(LiDAR)[ITC] (LiDAR) [leaf-off] (LiDAR) [leaf-off]	sri2_Range TIF_m_area	(Hyper.) [leaf-on] (LiDAR) [ITC]
<i>Mean DBH (cm)</i>	TIF_m_area varngwid_2	(LiDAR) [ITC] (LiDAR) [leaf-on]	TIF_m_area varngwid_2	(LiDAR) [ITC] (LiDAR) [leaf-on]
<i>STDev. of DBH (cm)</i>	TIF_m_area amp_p65	(LiDAR) [ITC] (LiDAR) [leaf-off]	amp_p65 TIF_m_area wd_p25_2	(LiDAR) [leaf-off] (LiDAR) [ITC] (LiDAR) [leaf-on]
<i>Basal area within 30x30m</i>	SI kurngamp	(Hyper.) [ITC] (LiDAR) [leaf-on]	meanht rendvi1_Max TIF_Space	(LiDAR) [leaf-off] (Hyper.) [leaf-off] (LiDAR) [ITC]
<i>Percentage big trees (DBH 40>x<80cm)</i>	mednght sri2_Min TIF_Space	(LiDAR) [leaf-off] (Hyper.) [leaf-on] (LiDAR) [ITC]	kurgamp_2 mednght TIF_Space	(LiDAR) [leaf-on] (LiDAR) [leaf-off] (LiDAR) [ITC]
<i>Percentage very big trees (DBH >80cm)</i>	kurwid lon_can_layer sri2_STD	(LiDAR) [leaf-off] (LiDAR) [leaf-on] (Hyper.) [leaf-on]	mrendvi1_STD no_tree	(Hyper.) [leaf-off] (Hyper.) [ITC]
<i>Percentage big trees (DBH >40cm)</i>	stdht_2 TIF_Space	(LiDAR) [leaf-on] (LiDAR) [ITC]	adevht_2 rgri2_Range TIF_m_area	(LiDAR) [leaf-on] (Hyper.) [leaf-on] (LiDAR) [ITC]
<i>Average height of the first live branch (m)</i>	loff_max_gap TIF_Space	(LiDAR) [leaf-off] (LiDAR) [ITC]	lon_max_gap stdngwid	(LiDAR) [leaf-on] (LiDAR) [leaf-off]
<i>STDev. of the height of the first live branch (m)</i>	loff_max_gap sri2_Max TIF_Space TIF_meanCV	(LiDAR) [leaf-off] (Hyper.) [leaf-on] (LiDAR) [ITC] (LiDAR) [ITC]	Domht TIF_Space TIF_to_area	(LiDAR) [leaf-off] (LiDAR) [ITC] (LiDAR) [ITC]
<i>Average crown horizontal area (m²)</i>	TIF_m_area mresri2_Max kurwid	(LiDAR) [ITC] (Hyper.) [leaf-on] (LiDAR) [leaf-off]	sipi1_Sum TIF_meanCV TIF_to_area	(Hyper.) [leaf-off] (LiDAR) [ITC] (LiDAR) [ITC]
<i>STDev. crown horizontal area (m²)</i>	TIF_m_area ari2_Min skwgwid mingamp_2	(LiDAR) [ITC] (Hyper.) [leaf-on] (LiDAR) [leaf-off] (LiDAR) [leaf-on]	ari2_Min TIF_m_area	(Hyper.) [leaf-on] (LiDAR) [ITC]
<i>Total crown horizontal area (m²)</i>	skwht_2 kurngwid_2	(LiDAR) [leaf-on] (LiDAR) [leaf-on]	kurwid_2 pri2_Max	(LiDAR) [leaf-on] (Hyper.) [leaf-on]

8.1.2 Forest composition within the plot

The associated R^2 values for the seven forest compositional metrics are given in Table 8.4 for both of the composite datasets. The first of the combinations was hyperspectral and DR LiDAR. For only two of these models did OLS methods produce higher R^2 values than AICc. For the second combination of hyperspectral and FW LiDAR, OLS produced three models with a higher R^2 value than the AICc counterpart.

Considering both modelling approaches, six of the total seven models for hyperspectral and DR had higher R^2 values than the combination of hyperspectral and FW (although the values were very similar for all but two field metrics).

The inputs to the regression models for a combination of hyperspectral and DR LiDAR are listed in Table 8.5. Table 8.6 documents the inputs for regression models for a combination of hyperspectral and FW LiDAR. A number of models use LiDAR or hyperspectral data, rather than a combination of both. For the first composite, hyperspectral and DR, 5 out of 7 models for OLS and 2 for AICc utilise DR LiDAR inputs only. Only one AICc model uses solely hyperspectral inputs. For the second composite, hyperspectral and FW, both OLS and AICc models include 4 which utilise solely FW LiDAR inputs. Only two AICc derived models utilise hyperspectral inputs only.

Table 8.4 – Forest composition metrics R^2 for hyperspectral and LiDAR combinations (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	<i>DR and Hyper.</i> OLS R^2	<i>DR and Hyper.</i> AICc R^2	<i>FW and Hyper.</i> OLS R^2	<i>FW and Hyper.</i> AICc R^2
<i>Number of native trees per 30x30m</i>	0.425	<u>0.642</u>	0.531	0.525
<i>Percentage of native trees</i>	<u>0.826</u>	0.779	0.599	0.729
<i>No. of tree species per 30x30m</i>	<u>0.775</u>	0.531	0.373	0.429
<i>Shannon index (by stem count)</i>	0.416	<u>0.662</u>	0.541	0.509
<i>Simpson index (by stem count)</i>	0.527	<u>0.587</u>	0.371	0.371
<i>Shannon index (by BA)</i>	0.419	<u>0.784</u>	0.757	0.592
<i>Simpson index (by BA)</i>	0.370	0.473	0.556	<u>0.570</u>

Table 8.5 – Hyperspectral and DR LiDAR metric inputs for each forest compositional metric regression model (Appendix B and C should be consulted for a description of remote sensing metric names.)

Metric Name:	DR and Hyper. OLSmodel input metrics		DR and Hyper. AICc model input metrics	
<i>Number of native trees per 30x30m</i>	SHA_Kur_2 p080	(LiDAR) [Leaf-on] (LiDAR) [Leaf-off]	sri2_Min vrei2_STD	(Hyper.) [Leaf-on] (Hyper.) [Leaf-on]
<i>Percentage of native trees</i>	TIF_TreeNo ari2_Min SHA_Ske arvi1_Min	(LiDAR) [ITC] (Hyper.) [Leaf-on] (LiDAR) [Leaf-off] (Hyper.) [Leaf-off]	arvi1_Min p060i_2 sri1_Sum	(Hyper.) [Leaf-off] (LiDAR) [Leaf-on] (Hyper.) [Leaf-off]
<i>No. of tree species per 30x30m</i>	SHN_Ske SIN_Var SIN_Ske_2 sipi1_Min	(LiDAR) [Leaf-off] (LiDAR) [Leaf-off] (LiDAR) [Leaf-on] (Hyper.) [Leaf-off]	p050 SIA_Var	(LiDAR) [Leaf-off] (LiDAR) [Leaf-off]
<i>Shannon index (by stem count)</i>	SIN_Kur_2 TIF_Space p050	(LiDAR) [Leaf-on] (LiDAR) [ITC] (LiDAR) [Leaf-off]	rgri1_Min rgri2_Min SHN_Ske SIA_Ske	(Hyper.) [Leaf-off] (Hyper.) [Leaf-on] (LiDAR) [Leaf-off] (LiDAR) [Leaf-off]
<i>Simpson index (by stem count)</i>	DEM_Slope SHN_STD_2	(LiDAR) [Leaf-off] (LiDAR) [Leaf-on]	rgri2_Min SHA_med_2 SIA_Kur_2	(Hyper.) [Leaf-on] (LiDAR) [Leaf-on] (LiDAR) [Leaf-on]
<i>Shannon index (by BA)</i>	DEM_Rough_2 TIF_m_area DEM_Slope_2	(LiDAR) [Leaf-on] (LiDAR) [ITC] (LiDAR) [Leaf-on]	ndvi2_STD no_nat_t SHA_Ske_2 SIA_Med sri2_Min	(Hyper.) [Leaf-on] (Hyper.) [ITC] (LiDAR) [Leaf-on] (LiDAR) [Leaf-off] (Hyper.) [Leaf-on]
<i>Simpson index (by BA)</i>	p075 SIN_Ske_2	(LiDAR) [Leaf-off] (LiDAR) [Leaf-on]	p070_2 SIN_Ske_2	(LiDAR) [Leaf-on] (LiDAR) [Leaf-on]

Table 8.6 – Hyperspectral and FW LiDAR metric inputs for each forest compositional metric regression model (Appendix B and C should be consulted for a description of remote sensing metric names)

Metric Name:	FW and Hyper. OLSmodel input metrics		FW and Hyper. AICc model input metrics	
<i>Number of native trees per 30x30m</i>	wd_p40 sri2_Min	(LiDAR) [Leaf-off] (Hyper.) [Leaf-on]	sri2_Min wd_p45	(Hyper.) [Leaf-on] (LiDAR) [Leaf-on]
<i>Percentage of native trees</i>	TIF_TreeNo ari2_Min	(LiDAR) [ITC] (Hyper.) [Leaf-on]	arvi1_Min sri1_Sum wbi2_STD	(Hyper.) [Leaf-off] (Hyper.) [Leaf-off] (Hyper.) [Leaf-on]
<i>No. of tree species per 30x30m</i>	wd_p30_2 stdgwid_2	(LiDAR) [Leaf-on] (LiDAR) [Leaf-on]	rendvi1_Max rendvi1_STD sipi1_Range	(Hyper.) [Leaf-off] (Hyper.) [Leaf-off] (Hyper.) [Leaf-off]
<i>Shannon index (by stem count)</i>	skwngamp_2 sri2_Max vrei2_Min	(LiDAR) [Leaf-on] (Hyper.) [Leaf-on] (Hyper.) [Leaf-on]	skwngamp_2 TIF_to_area	(LiDAR) [Leaf-on] (LiDAR) [ITC]
<i>Simpson index (by stem count)</i>	kurngamp_2 varht_2	(LiDAR) [Leaf-on] (LiDAR) [Leaf-on]	kurngamp_2 varht_2	(LiDAR) [Leaf-on] (LiDAR) [Leaf-on]
<i>Shannon index (by BA)</i>	vargwid_2 TIF_m_area ht_p15_2 wd_p15_2	(LiDAR) [Leaf-on] (LiDAR) [ITC] (LiDAR) [Leaf-on] (LiDAR) [Leaf-on]	maxgwid stdnght_2 varamp_2	(LiDAR) [Leaf-off] (LiDAR) [Leaf-on] (LiDAR) [Leaf-on]
<i>Simpson index (by BA)</i>	ht_p80 medngamp_2	(LiDAR) [Leaf-off] (LiDAR) [Leaf-on]	medamp_2 stdht_2	(LiDAR) [Leaf-on] (LiDAR) [Leaf-on]

8.1.3 Deadwood within the plot

The R^2 values contained within Table 8.7 relate to the four deadwood metrics regression models produced from both composite remote sensing datasets. The first of the combinations was hyperspectral and DR LiDAR. For only two of four models OLS methods produced higher R^2 values. The highest R^2 values ranged from 0.598 to 0.902 for the deadwood metrics. For the second combination of hyperspectral and FW LiDAR, OLS produced three out of four models with a higher R^2 value than the AICc counterpart. The highest R^2 values ranged from 0.639 to 0.955 for the deadwood metrics.

In comparing the R^2 values between the two composite models, all four of the best models for hyperspectral and FW had higher R^2 values than the hyperspectral and DR LiDAR composite.

The inputs to the regression models for the combination of hyperspectral and DR LiDAR and the combination of hyperspectral and FW LiDAR are listed in Table 8.8 and Table 8.9 respectively. Measures of overstorey species diversity are used to predict standing deadwood decay class for 3 of 4 regression models. Hyperspectral VI information is used in the majority of models. A number of models did not use in a combination of LiDAR and hyperspectral metrics. For both composites, one OLS model uses only LiDAR metrics as input, whereas two AICc models utilise solely hyperspectral inputs.

Table 8.7 – Forest deadwood metrics R^2 for hyperspectral and LiDAR combinations (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	DR and Hyper. OLS R^2	DR and Hyper. AICc R^2	FW and Hyper. OLS R^2	FW and Hyper. AICc R^2
<i>Standing deadwood volume (m^3)</i>	0.748	<u>0.881</u>	<u>0.919</u>	0.833
<i>Standing deadwood decay class</i>	0.476	<u>0.902</u>	0.955	<u>0.955</u>
<i>Downed deadwood volume (m^3)</i>	<u>0.598</u>	0.485	<u>0.639</u>	0.485
<i>Downed deadwood decay class</i>	<u>0.681</u>	0.580	<u>0.750</u>	0.717

Table 8.8 – Hyperspectral and DR LiDAR metric inputs for each forest deadwood metric regression model (Appendix B and C should be consulted for a description of remote sensing metric names.)

Metric Name:	DR and Hyper. OLSmodel input metrics	DR and Hyper. AICc model input metrics
<i>Standing deadwood volume (m³)</i>	SIN_Ske (LiDAR) [Leaf-off] ari2_Max (Hyper.) [Leaf-off]	p080i (LiDAR) [Leaf-off] mrendvi1_STD (Hyper.) [Leaf-off] rgri2_STD (Hyper.) [Leaf-off]
<i>Standing deadwood decay class</i>	vrei1_Range (Hyper.) [Leaf-off]	mresri1_Mean (Hyper.) [Leaf-off] SI (Hyper.) [ITC] sri2_Max (Hyper.) [Leaf-on]
<i>Downed deadwood volume (m³)</i>	mresri1_Max (Hyper.) [Leaf-off] ari2_Max (Hyper.) [Leaf-on] SHN_Med (LiDAR) [Leaf-off]	ari2_Max (Hyper.) [Leaf-on] mresri1_Max (Hyper.) [Leaf-off]
<i>Downed deadwood decay class</i>	SIN_Med_2 (LiDAR) [Leaf-on] DEM_Slope_2 (LiDAR) [Leaf-on]	mrendvi2_STD (Hyper.) [Leaf-on] SIN_mean_2 (LiDAR) [Leaf-on]

Table 8.9 – Hyperspectral and FW LiDAR metric inputs for each forest deadwood metric regression model (Appendix B and C should be consulted for a description of remote sensing metric names.)

Metric Name:	FW and Hyper. OLSmodel input metrics	FW and Hyper. AICc model input metrics
<i>Standing deadwood volume (m³)</i>	TIF_STD (LiDAR) [ITC] wd_p25_2 (LiDAR) [Leaf-on] skwngamp (LiDAR) [Leaf-off]	rgri2_STD (Hyper.) [Leaf-on] vrei1_STD (Hyper.) [Leaf-off]
<i>Standing deadwood decay class</i>	sri1_Sum (Hyper.) [Leaf-off] sri2_Max (Hyper.) [Leaf-off] SH (Hyper.) [ITC] skwgid (LiDAR) [Leaf-off]	SH (Hyper.) [ITC] skwgid (LiDAR) [Leaf-off] sri1_Sum (Hyper.) [Leaf-off] sri2_Max (Hyper.) [Leaf-on]
<i>Downed deadwood volume (m³)</i>	rendvi1_Max (Hyper.) [Leaf-off] medht (LiDAR) [Leaf-off] ari2_Range (Hyper.) [Leaf-on]	ari2_Max (Hyper.) [Leaf-on] mresri1_Max (Hyper.) [Leaf-off]
<i>Downed deadwood decay class</i>	varngamp_2 (LiDAR) [Leaf-on] skwgid_2 (LiDAR) [Leaf-on] wd_p15_2 (LiDAR) [Leaf-on] medngamp_2 (LiDAR) [Leaf-on]	mrendvi2_STD (Hyper.) [Leaf-on] skwid_2 (LiDAR) [Leaf-on] varngamp_2 (LiDAR) [Leaf-on]

8.1.4 Understorey metrics within the plot

There were 12 forest understorey metrics which are summarised in Table 8.10 along with the associated R^2 values for the regression models applied to the two composite datasets. The first of the combinations was hyperspectral and DR LiDAR. For only five of twelve models OLS methods produced higher R^2 values. The highest R^2 values ranged from 0.538 to 0.951 for the understorey metrics. For the second combination of hyperspectral and FW LiDAR, the number of vascular species could not be estimated using the AICc method. OLS produced three out of twelve models with a higher R^2 value than the AICc counterpart. The highest R^2 values ranged from 0.404 to 0.973 for the deadwood metrics.

In comparing the R^2 values between the two composite models, three of the total twelve models for hyperspectral and FW had higher R^2 values compared with hyperspectral and DR LiDAR.

The inputs to the regression models for the combination of hyperspectral and DR LiDAR are listed in Table 8.11. Table 8.12 documents the inputs for the regression models for the combination of hyperspectral and FW LiDAR. Many of the regression model inputs utilise a combination of hyperspectral and LiDAR inputs, for both composite datasets. The first composite, DR and hyperspectral, had four OLS models and six AICc models which utilized DR metric only, the remainder were all combinations of hyperspectral and DR LiDAR. The second composite dataset, FW and hyperspectral, had two OLS and 5 AICc models which utilised only FW LiDAR inputs, in addition to 4 OLS and 1 AICc models which used only hyperspectral metrics as input.

Table 8.10 – Forest understorey metrics R^2 for hyperspectral and LiDAR combinations (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	<i>DR and Hyper.</i> OLS R^2	<i>DR and Hyper.</i> AICc R^2	<i>FW and Hyper.</i> OLS R^2	<i>FW and Hyper.</i> AICc R^2
<i>No. saplings per 30x30m</i>	0.642	<u>0.922</u>	0.535	<u>0.973</u>
<i>No. native saplings per 30x30m</i>	0.591	<u>0.896</u>	0.613	<u>0.968</u>
<i>Percentage of native saplings per 30x30m</i>	0.692	<u>0.951</u>	0.692	<u>0.894</u>
<i>No. of sapling species per 30x30m</i>	0.494	<u>0.922</u>	0.848	<u>0.895</u>
<i>No. seedlings per 30x30m</i>	<u>0.609</u>	0.527	0.498	<u>0.568</u>
<i>No. native seedlings per 30x30m</i>	0.430	<u>0.538</u>	0.779	<u>0.779</u>
<i>Percentage of native seedlings per 30x30m</i>	<u>0.763</u>	0.577	0.605	<u>0.865</u>
<i>No. of seedling species per 30x30m</i>	<u>0.740</u>	0.553	<u>0.860</u>	0.532
<i>Seedlings Shannon Index for native species</i>	0.674	<u>0.789</u>	<u>0.792</u>	0.746
<i>No. vascular species</i>	<u>0.674</u>	0.547	0.404	-
<i>No. bryophyte species</i>	<u>0.737</u>	0.625	0.687	0.687
<i>Percentage of bare soil within 30x30m</i>	0.761	<u>0.793</u>	<u>0.840</u>	0.749

Table 8.11 – Hyperspectral and DR LiDAR metric inputs for each forest understorey metric regression model (Appendix B and C should be consulted for a description of remote sensing metric names.)

Metric Name:	DR and Hyper. OLS model input metrics	DR and Hyper. AICc model input metrics
<i>No. saplings per 30x30m</i>	mresri2_Max (Hyper.) [Leaf-on] SIG_Ske (LiDAR) [Leaf-off]	p060i_2 (LiDAR) [Leaf-on] p060_2 (LiDAR) [Leaf-on] SIA_Mean (LiDAR) [Leaf-on]
<i>No. native saplings per 30x30m</i>	mresri2_Max (Hyper.) [Leaf-on] SIG_Ske (LiDAR) [Leaf-off]	p060i_2 (LiDAR) [Leaf-on] SHN_Med_2 (LiDAR) [Leaf-on] SIA_Mean (LiDAR) [Leaf-off]
<i>Percentage of native saplings per 30x30m</i>	mresri2_STD (Hyper.) [Leaf-on]	SHN_Ske_2 (LiDAR) [Leaf-on] SIN_Mean_2 (LiDAR) [Leaf-on] SIN_Var (LiDAR) [Leaf-off]
<i>No. of sapling species per 30x30m</i>	sri2_Max (Hyper.) [Leaf-on] SIA_Ske (LiDAR) [Leaf-off]	pri1_Max (Hyper.) [Leaf-off] SHA_mean_2 (LiDAR) [Leaf-on]
<i>No. seedlings per 30x30m</i>	SIG_Kur (LiDAR) [Leaf-off] p060i_2 (LiDAR) [Leaf-on] TIF_meanCV (LiDAR) [ITC]	SIG_MAD (LiDAR) [Leaf-off] SIG_Mean_2 (LiDAR) [Leaf-on]
<i>No. native seedlings per 30x30m</i>	ari2_Range (Hyper.) [Leaf-on] SIG_Var (LiDAR) [Leaf-off]	spec_no (Hyper.) [ITC] TIF_meanCV (LiDAR) [ITC]
<i>Percentage of native seedlings per 30x30m</i>	p065i_2 (LiDAR) [Leaf-on] SIA_Med_2 (LiDAR) [Leaf-on] ari2_Sum (Hyper.) [Leaf-on]	ari2_Min (Hyper.) [Leaf-on] mresri2_Max (Hyper.) [Leaf-on] SIA_Mean_2 (LiDAR) [Leaf-on]
<i>No. of seedling species per 30x30m</i>	mresri2_Max (Hyper.) [Leaf-on] rgri1_Min (Hyper.) [Leaf-off] DEM_Slope (LiDAR) [Leaf-off]	p060i_2 (LiDAR) [Leaf-on] SIA_Ske_2 (LiDAR) [Leaf-on]
<i>Seedlings Shannon index for native species</i>	PAIP_2 (LiDAR) [Leaf-on]	p045_2 (LiDAR) [Leaf-on] rgri1_Min (Hyper.) [Leaf-off] SHN_Med_2 (LiDAR) [Leaf-on] spec_no (Hyper.) [ITC]
<i>No. vascular species</i>	sri2_Range (Hyper.) [Leaf-on] sri2_STD (Hyper.) [Leaf-on] p030i_2 (LiDAR) [Leaf-on]	p025i_2 (LiDAR) [Leaf-on] SHN_MAD_2 (LiDAR) [Leaf-on] sri2_Min (Hyper.) [Leaf-on]
<i>No. bryophyte species</i>	p085i (LiDAR) [Leaf-off] SHN_Med (LiDAR) [Leaf-off] lon_can_layers (LiDAR) [Leaf-on]	p010_2 (LiDAR) [Leaf-on] rgri2_Range (Hyper.) [Leaf-on] SIN_Kur (LiDAR) [Leaf-off]
<i>Percentage of bare soil within 30x30m</i>	p020i_2 (LiDAR) [Leaf-on]	p010i_2 (LiDAR) [Leaf-on] p025i_2 (LiDAR) [Leaf-on]

Table 8.12 – Hyperspectral and FW LiDAR metric inputs for each forest understorey metric regression model (Appendix B and C should be consulted for a description of remote sensing metric names.)

Metric Name:	FW and Hyper. OLS model input metrics	FW and Hyper. AICc model input metrics
<i>No. saplings per 30x30m</i>	mresri2_Max (Hyper.) [Leaf-on] vrei2_Range (Hyper.) [Leaf-on]	ht_p35_2 (LiDAR) [Leaf-on] medht (LiDAR) [Leaf-off] wd_p35_2 (LiDAR) [Leaf-on]
<i>No. native saplings per 30x30m</i>	mresri2_Range (Hyper.) [Leaf-on] vrei2_Range (Hyper.) [Leaf-on] wd_p35_2 (LiDAR) [Leaf-on]	ht_p35_2 (LiDAR) [Leaf-on] medht (LiDAR) [Leaf-off] wd_p35_2 (LiDAR) [Leaf-on]
<i>Percentage of native saplings per 30x30m</i>	mresri2_STD (Hyper.) [Leaf-on]	ari2_Min (Hyper.) [Leaf-on] wd_p45 (LiDAR) [Leaf-on]
<i>No. of sapling species per 30x30m</i>	ht_p35_2 (LiDAR) [Leaf-on] varamp (LiDAR) [Leaf-off] mresri2_Max (Hyper.) [Leaf-on]	ari2_Min (Hyper.) [Leaf-on] ndvi2_Min (Hyper.) [Leaf-on] no_nat_t (Hyper.) [ITC]
<i>No. seedlings per 30x30m</i>	varnght (LiDAR) [Leaf-off] spec_no (Hyper.) [ITC]	medht (LiDAR) [Leaf-off] mednght (LiDAR) [Leaf-off] nat_spec (Hyper.) [ITC]
<i>No. native seedlings per 30x30m</i>	kurngamp (LiDAR) [Leaf-off] spec_no (Hyper.) [ITC] vargwid (LiDAR) [Leaf-off] TIF_meanCV (LiDAR) [ITC]	kurngamp (LiDAR) [Leaf-off] spec_no (Hyper.) [ITC] TIF_meanCV (LiDAR) [ITC] vargwid (LiDAR) [Leaf-off]
<i>Percentage of native seedlings per 30x30m</i>	medgamp_2 (LiDAR) [Leaf-on]	medgamp_2 (LiDAR) [Leaf-on]
<i>No. of seedling species per 30x30m</i>	mresri2_Max (Hyper.) [Leaf-on] rgri1_Min (Hyper.) [Leaf-off] spec_no (Hyper.) [ITC] vrei2_STD (Hyper.) [Leaf-on]	sri2_Max (Hyper.) [Leaf-on] stdgwid (LiDAR) [Leaf-off]
<i>Seedlings Shannon index for native species</i>	pri2_Range (Hyper.) [Leaf-on] skwnght (LiDAR) [Leaf-off] skwgwid_2 (LiDAR) [Leaf-on] rgri1_Min (Hyper.) [Leaf-off]	adevht_2 (LiDAR) [Leaf-on] spec_no (Hyper.) [ITC]
<i>No. vascular species</i>	mresri2_Max (Hyper.) [Leaf-on] mresri1_STD (Hyper.) [Leaf-off]	-
<i>No. bryophyte species</i>	skwngamp (LiDAR) [Leaf-off] lon_can_layer (LiDAR) [Leaf-on]	lon_can_layer (LiDAR) [Leaf-on] skwngamp (LiDAR) [Leaf-off]
<i>Percentage of bare soil within 30x30m</i>	kurgamp_2 (LiDAR) [Leaf-on] no_tree (Hyper.) [ITC] spec_no (Hyper.) [ITC]	ht_p15_2 (LiDAR) [Leaf-on] loff_can_depth (LiDAR) [Leaf-off] meannght_2 (LiDAR) [Leaf-on]

8.1.5 Metrics required for condition index construction

OLS regression analysis using either remote sensing data combination could not be completed for these two remaining metrics, the models failing to achieve significance. The two forest structural metrics and its associated R^2 values for AICc models are listed in Table 8.13. The first of the combinations was hyperspectral and DR LiDAR, where both models returned high R^2 values, i.e. above 0.7. Likewise the combination of hyperspectral and FW LiDAR returned high R^2 values, i.e. above 0.7.

In comparing the R^2 values between the two composite models, similar R^2 values were recorded, i.e. within 0.09 for both models.

Table 8.14 documents the input metrics to the regression models for both the combination of hyperspectral and DR LiDAR and the combination of hyperspectral and FW LiDAR. All model inputs are composed of both hyperspectral and LiDAR data, except for one, the DR and hyperspectral composite for estimating the index of vertical separation.

Table 8.13 – Condition index construction metrics R^2 for hyperspectral and LiDAR combinations (all models $p < 0.05$). Underlined values indicate highest R^2 value.

Metric Name:	DR and Hyper. AICc R^2	FW and Hyper. AICc R^2
<i>Sum of the DBH differences between neighbouring trees (cm)</i>	<u>0.950</u>	0.864
<i>Index of vertical separation</i>	0.730	<u>0.736</u>

Table 8.14 – Combined remote sensing metric inputs for each forest condition index metric regression model (Appendix B and C should be consulted for a description of remote sensing metric names)

Metric Name:	DR and Hyper. AICc model input metrics		FW and Hyper. AICc model input metrics	
<i>Sum of the DBH differences between neighbouring trees (cm)</i>	cri1_Sum	(Hyper.) [Leaf-off]	kurht_2	(LiDAR) [Leaf-on]
	SHA_Kur_2	(LiDAR) [Leaf-on]	no_tree	(Hyper.) [ITC]
	sri2_Min	(Hyper.) [Leaf-on]		
	TIF_m_area	(LiDAR) [ITC]		
	TIF_meanCV	(LiDAR) [ITC]		
<i>Index of vertical separation</i>	CC	(LiDAR) [Leaf-off]	amp_p15	(LiDAR) [Leaf-off]
	p075	(LiDAR) [Leaf-off]	rendvi2_Min	(Hyper.) [Leaf-on]
	TIF_to_area	(LiDAR) [ITC]	stdht	(LiDAR) [Leaf-off]
			TIF_to_area	(LiDAR) [ITC]

8.2 Validation of remote sensing derived field-level metrics via comparison with field measurements

This results section outlines a summary of the various statistical model estimates derived from remote sensing and their relation to validation field measurements.

8.2.1 Hyperspectral and DR LiDAR derived statistical models

The results of the regression models from the combined hyperspectral and DR LiDAR datasets were compared with the 2012 field data measurements (see Table 8.15). RMSE and NRMSE were calculated for each of the remote sensing derived models. Five out of twenty-one of the model pairs for overstorey composition recorded NRMSE values above 0.5. The field-level metrics which could not be modelled accurately were: the average tree NN spacing, the standard deviation of tree NN spacing, the mean DBH, the percentage of very big trees (DBH>80cm), and the SH index (by basal area). The OLS model for predicting basal area and the standard deviation of the height of the first live branch also had a high NRMSE value. RMSE values for OLS models were lower for thirteen out of twenty-one models, although often the values are very similar.

Table 8.16 reports the RMSE and NRMSE values for understorey field-level metrics. Four of the field-level metrics produced values of NRMSE above 0.5 for both OLS and AICc models or could not be modelled. The field-level metrics which could not be modelled were: the percentage of native saplings, the number of sapling species, the number of bryophyte species and the index of vertical separation. NRMSE values for AICc models of standing deadwood volume and the percentage bare soil were above 0.5, as were OLS models for DDW volume.

Table 8.15 – The OLS and AICc hyperspectral and DR LiDAR regression models. Part 1: overstorey composition. Underlined values denote the lowest RMSE value for each field metric.

Field metric	OLS RMSE	OLS NRMSE	AICc RMSE	AICc NRMSE
<i>Number of trees per 30x30m</i>	<u>20.211</u>	<u>0.273</u>	20.001	0.270
<i>Number of native trees per 30x30m</i>	<u>15.141</u>	<u>0.291</u>	15.676	0.301
<i>Percentage of native trees</i>	37.264	0.373	<u>33.812</u>	<u>0.338</u>
<i>No. of tree species per 30x30m</i>	<u>1.503</u>	<u>0.250</u>	1.542	0.257
<i>Average tree NN spacing (m)</i>	1.802	0.542	1.919	0.578
<i>STDev. of tree NN spacing (m)</i>	0.960	1.155	1.101	1.325
<i>Mean DBH (cm)</i>	15.151	0.520	15.406	0.529
<i>STDev. of DBH (cm)</i>	16.455	0.337	<u>14.976</u>	<u>0.307</u>
<i>Basal area within 30x30m (m²)</i>	1.786	0.557	<u>1.454</u>	<u>0.453</u>
<i>(i) Percentage big trees(DBH 40>x<80cm)</i>	<u>22.992</u>	<u>0.328</u>	25.048	0.358
<i>(ii) Percentage very big trees (DBH >80cm)</i>	10.294	1.698	12.618	2.082
<i>(iii) Percentage big trees (DBH >40cm)</i>	<u>26.218</u>	<u>0.375</u>	27.052	0.386
<i>Shannon index (by stem count)</i>	<u>0.442</u>	<u>0.308</u>	0.654	0.456
<i>Simpson index (by stem count)</i>	0.399	0.309	<u>0.321</u>	<u>0.249</u>
<i>Shannon index (by BA)</i>	0.577	0.751	0.671	0.873
<i>Simpson index (by BA)</i>	<u>0.626</u>	<u>0.217</u>	0.649	0.225
<i>Average height of the first live branch (m)</i>	2.879	0.266	<u>2.658</u>	<u>0.245</u>
<i>STDev. of the height of the first live branch (m)</i>	2.790	0.541	<u>2.232</u>	<u>0.433</u>
<i>Average crown horizontal area (m²)</i>	35.022	0.402	<u>28.514</u>	<u>0.327</u>
<i>STDev. crown horizontal area (m²)</i>	26.788	0.468	<u>17.766</u>	<u>0.311</u>
<i>Total crown horizontal area (m²)</i>	<u>510.517</u>	<u>0.354</u>	641.159	0.444

Table 8.16 – The OLS and AICc for hyperspectral and DR LiDAR regression models. Part 2: understorey composition. Underlined values denote the lowest RMSE value for each field metric.

Field metric	OLS RMSE	OLS NRMSE	AICc RMSE	AICc NRMSE
<i>Standing deadwood volume (m³)</i>	2.218	0.268	<u>4.508</u>	<u>0.545</u>
<i>Standing deadwood decay class</i>	0.237	0.355	<u>0.217</u>	<u>0.325</u>
<i>Downed deadwood volume (m³)</i>	5.111	0.562	<u>4.382</u>	<u>0.482</u>
<i>Downed deadwood decay class</i>	0.216	0.460	<u>0.193</u>	<u>0.412</u>
<i>No. saplings per 30x30m</i>	24.791	0.230	<u>24.406</u>	<u>0.226</u>
<i>No. native saplings per 30x30m</i>	<u>5.731</u>	<u>0.151</u>	12.350	0.325
<i>Percentage of native saplings per 30x30m</i>	78.581	0.786	60.853	0.609
<i>No. of sapling species per 30x30m</i>	1.658	0.553	1.533	0.511
<i>No. seedlings per 30x30m</i>	450.132	0.481	<u>450.051</u>	<u>0.481</u>
<i>No. native seedlings per 30x30m</i>	256.432	0.356	<u>234.340</u>	<u>0.325</u>
<i>Percentage of native seedlings per 30x30m</i>	<u>42.172</u>	<u>0.422</u>	44.803	0.448
<i>No. of seedling species per 30x30m</i>	4.099	0.315	<u>3.391</u>	<u>0.261</u>
<i>Seedlings Shannon index for native species</i>	<u>0.766</u>	<u>0.475</u>	0.780	0.483
<i>No. vascular species</i>	2.793	0.465	<u>1.987</u>	<u>0.331</u>
<i>No. bryophyte species</i>	7.812	0.781	7.040	0.704
<i>Percentage of bare soil within 30x30m</i>	<u>46.340</u>	<u>0.488</u>	48.202	0.507
<i>Sum of the DBH differences between neighbouring trees (cm)</i>	-	-	1.693	0.057
<i>Index of vertical separation</i>	-	-	0.327	0.649

8.2.2 Hyperspectral and FW LiDAR derived statistical models

The results of the regression models from the combined hyperspectral and FW LiDAR datasets for deriving forest overstorey metrics were compared with the 2012 field data measurements (see Table 8.17). RMSE and NRMSE were calculated for each of the remote sensing derived models. Four model pairs recorded NRMSE values above 0.5, these were: the average tree NN spacing, the standard deviation of tree NN spacing, the percentage of very big trees (DBH>80cm) and the SH index (by basal area). OLS estimates for mean DBH also had NRMSE values above 0.5. AICc estimates of total basal area and the standard deviation of crown horizontal area also had NRSME values above 0.5. OLS models had lower RMSE values for a total of six out of the remaining thirteen metrics. For all but one of these, RMSE values were similar.

Table 8.18 reports the RMSE and NRMSE values for the forest understorey metrics. Two out of a total of eighteen model pairs reported NRMSE values of greater than 0.5, these were the percentage of native saplings and the number of bryophyte species. AICc estimates for standing deadwood reported NRSME values higher than 0.5, whereas four OSL model estimates reported NRMSE values above 0.5, DDW volume, DDW class, seedlings SH index and the percentage of bare soil.

Table 8.17 – The OLS and AICc hyperspectral and FW LiDAR regression models. Part 1: overstorey composition. Underlined values denote the lowest RMSE value for each field metric.

Field metric	OLS RMSE	OLS NRMSE	AICc RMSE	AICc NRMSE
<i>Number of trees per 30x30m</i>	<u>19.138</u>	<u>0.259</u>	20.786	0.281
<i>Number of native trees per 30x30m</i>	13.800	0.265	<u>12.637</u>	<u>0.243</u>
<i>Percentage of native trees</i>	<u>31.820</u>	<u>0.318</u>	36.500	0.365
<i>No. of tree species per 30x30m</i>	<u>2.145</u>	<u>0.357</u>	2.387	0.398
<i>Average tree NN spacing (m)</i>	1.802	0.542	1.825	0.549
<i>STDev. of tree NN spacing (m)</i>	0.431	0.518	0.880	1.058
<i>Mean DBH (cm)</i>	15.406	0.529	<u>12.627</u>	<u>0.433</u>
<i>STDev. of DBH (cm)</i>	9.458	0.194	<u>9.100</u>	<u>0.187</u>
<i>Basal area within 30x30m (m²)</i>	<u>1.150</u>	<u>0.359</u>	1.734	0.541
<i>(i) Percentage big trees (DBH 40 > x < 80cm)</i>	25.530	0.365	<u>19.930</u>	<u>0.285</u>
<i>(ii) Percentage very big trees (DBH > 80cm)</i>	11.868	1.958	10.330	1.704
<i>(iii) Percentage big trees (DBH > 40cm)</i>	<u>26.410</u>	<u>0.377</u>	31.555	0.451
<i>Shannon index (by stem count)</i>	0.523	0.365	0.578	0.404
<i>Simpson index (by stem count)</i>	0.287	0.222	0.287	0.222
<i>Shannon index (by BA)</i>	0.575	0.748	0.499	0.649
<i>Simpson index (by BA)</i>	0.856	0.296	<u>0.795</u>	<u>0.275</u>
<i>Average height of the first live branch (m)</i>	<u>2.657</u>	<u>0.245</u>	4.896	0.452
<i>STDev. of the height of the first live branch (m)</i>	2.206	0.428	<u>1.871</u>	<u>0.363</u>
<i>Average crown horizontal area (m²)</i>	32.776	0.376	<u>31.829</u>	<u>0.366</u>
<i>STDev. crown horizontal area (m²)</i>	<u>27.172</u>	<u>0.475</u>	29.093	0.509
<i>Total crown horizontal area (m²)</i>	682.213	0.473	<u>581.459</u>	<u>0.403</u>

Table 8.18 – The OLS and AICc for hyperspectral and FW LiDAR regression models. Part 2: understorey composition. Underlined values denote the lowest RMSE value for each field metric.

Field metric	OLS RMSE	OLS NRMSE	AICc RMSE	AICc NRMSE
<i>Standing deadwood volume (m³)</i>	<u>1.360</u>	<u>0.165</u>	4.296	0.520
<i>Standing deadwood decay class</i>	0.224	0.336	<u>0.224</u>	<u>0.336</u>
<i>Downed deadwood volume (m³)</i>	5.431	0.597	<u>4.395</u>	<u>0.483</u>
<i>Downed deadwood decay class</i>	0.250	0.532	<u>0.227</u>	<u>0.483</u>
<i>No. saplings per 30x30m</i>	<u>25.189</u>	<u>0.233</u>	26.502	0.245
<i>No. native saplings per 30x30m</i>	<u>9.168</u>	<u>0.241</u>	11.465	0.302
<i>Percentage of native saplings per 30x30m</i>	78.581	0.786	71.081	0.711
<i>No. of sapling species per 30x30m</i>	1.072	0.357	1.025	0.342
<i>No. seedlings per 30x30m</i>	<u>260.460</u>	<u>0.278</u>	309.812	0.331
<i>No. native seedlings per 30x30m</i>	264.820	0.368	264.820	0.368
<i>Percentage of native seedlings per 30x30m</i>	40.700	0.407	40.700	0.407
<i>No. of seedling species per 30x30m</i>	4.658	0.358	<u>3.090</u>	<u>0.238</u>
<i>Seedlings Shannon index for native species</i>	0.912	0.565	<u>0.703</u>	<u>0.436</u>
<i>No. vascular species</i>	2.890	0.482	-	-
<i>No. bryophyte species</i>	11.088	1.109	11.088	1.109
<i>Percentage of bare soil within 30x30m</i>	49.276	0.519	<u>39.955</u>	<u>0.421</u>
<i>Sum of the DBH differences between neighbouring trees (cm)</i>	-	-	2.806	0.094
<i>Index of vertical separation</i>	-	-	0.201	0.397

8.3A review of all regression models created from the five remote sensing datasets

This section consists of a listing of the statistical models applied to the five remote sensing datasets in order to estimate field plot-level metrics. These were: (i) hyperspectral, (ii) DR LiDAR, (iii) FW LiDAR, (iv) combined hyperspectral and DR LiDAR, and (v) combined hyperspectral and FW LiDAR. The following overviews the results presented in sections 6.3, 7.3, 7.4 and 8.1 to highlight the comparative results. The R^2 value is used to describe how well a regression line fits a set of data, in this case the initial field data collected in 2010. Each of the models presented were statistically significant ($p < 0.05$).

Each of the 39 field metrics had a total of ten statistical models calculated in order to predict the variable from the five different combinations of remote sensing metrics using the two regression approaches (OLS and AICc). The model results will be presented by field metric type: (i) forest structure within plot; (ii) forest composition within plot; (iii) deadwood within plot; and (iv) regeneration and understorey structure and composition.

8.3.1 Forest structure within plot

A listing of all 14 forest structural metrics and their associated R^2 values are listed in Table 8.19. When considering the statistical methods, OLS and AICc, very similar R^2 values are often produced with a typical variation of ± 0.12 and a range of variation of 0.31. On average, AICc models produce higher R^2 values in comparison with OLS models. The largest differences are encountered for the prediction of plot basal area for all RS inputs, except hyperspectral combined with FW LiDAR. In addition, there was a difference in R^2 between DR LiDAR OLS and AICc values for the percentage of very big trees (DBH > 80cm).

The predictive power of different remote sensing inputs for structural metrics does vary, most notably between hyperspectral and LiDAR datasets. A number of models derived from hyperspectral data are comparable with the LiDAR and composite equivalents, such as the number of trees per 30x30m plot and the percentage of very big trees (DBH > 80cm). A comparison of DR and FW LiDAR model values against composite model values generally yields little difference. Composite datasets models relating to the plot metrics of basal area and percentage of very big trees (DBH > 80cm) had higher R^2 values than for the hyperspectral, DR or FW equivalents.

In review of the total 14 forest structural metrics, five were modelled with the highest R^2 values using the hyperspectral and DR LiDAR composite using the AICc approach, three using the FW LiDAR data only using the OLS approach, and the remaining six are modelled with high R^2 values used different remote sensing dataset combinations and modelling approaches. Only the DR LiDAR data with an AICc modelling approach and hyperspectral and DR LiDAR composite with an OLS modelling approach did not generate a best fit model.

Table 8.19 – Forest structural metric R^2 values for all datasets (all models $p < 0.05$). Blue indicates the highest R^2 value for each row.

Indicator Metric	Hyper. OLS (R^2)	Hyper. AICc (R^2)	DR OLS (R^2)	DR AICc (R^2)	FW OLS (R^2)	FW AICc (R^2)	H+DR OLS (R^2)	H+DR AICc (R^2)	H+FW OLS (R^2)	H+FW/AI Cc (R^2)
<i>Number of trees per 30x30m</i>	0.749	0.808	0.862	0.894	0.897	0.672	0.804	0.912	0.699	0.771
<i>Average tree spacing (m)</i>	0.482	0.609	0.799	0.794	0.786	0.906	0.785	0.854	0.785	0.819
<i>STDev. Of tree spacing (m)</i>	0.571	0.812	0.753	0.752	0.849	0.682	0.771	0.771	0.835	0.664
<i>Mean DBH (cm)</i>	0.549	0.691	0.663	0.877	0.933	0.796	0.664	0.755	0.796	0.796
<i>STDev. of DBH (cm)</i>	0.768	0.796	0.625	0.739	0.823	0.819	0.779	0.867	0.692	0.819
<i>Basal area within 30x30m</i>	0.753	0.484	0.580	0.462	0.691	0.658	0.761	0.537	0.573	0.822
<i>Percentage big trees (DBH 40>x<80cm)</i>	0.505	0.554	0.628	0.901	0.913	0.651	0.850	0.925	0.809	0.915
<i>Percentage very big trees (DBH >80cm)</i>	0.771	0.744	0.628	0.391	0.642	0.471	0.654	0.799	0.592	0.744
<i>Percentage big trees (DBH >40cm)</i>	0.735	0.581	0.839	0.942	0.939	0.920	0.924	0.942	0.915	0.908
<i>Average height of the first live branch (m)</i>	0.692	0.672	0.896	0.849	0.947	0.876	0.820	0.876	0.876	0.805
<i>STDev. of the height of the first live branch (m)</i>	0.464	0.588	0.824	0.712	0.509	0.817	0.808	0.629	0.743	0.805
<i>Average crown horizontal area (m²)</i>	0.564	0.880	0.748	0.864	0.862	0.629	0.852	0.859	0.86	0.828
<i>STDev. Crown horizontal area (m²)</i>	0.520	0.727	0.857	0.818	0.811	0.633	0.818	0.727	0.892	0.796
<i>Total Crown horizontal area (m²)</i>	0.820	0.716	0.748	0.731	0.521	0.690	0.601	0.662	0.605	0.761

Cells highlighted in blue denote the highest R^2 value for each field metric.

8.3.2 Forest composition within plot

The two regression approaches were applied to the seven forest compositional metrics and the R^2 values for each of the derived models are given in Table 8.20. A number of significant models could not be obtained for compositional metrics when using hyperspectral and DR LiDAR datasets; FW LiDAR and composite models did not have this issue. Hyperspectral models failed to predict the Simpson index (weighted by basal area) with both methods, and OLS methods failed to predict the number of tree species per plot extent and the Shannon index (weighted by stem counts). DR LiDAR OLS models failed to predict the Simpson (weighted by stem counts) and Shannon (weighted by basal area) indices of diversity.

When considering the statistical methods, generally AICc methods deliver higher R^2 values. However the average difference is small at ± 0.11 , as illustrated in Table 8.20, with a range of variation of 0.43. The largest differences in R^2 values are found when considering the Simpson and Shannon diversity indices. For hyperspectral data, the Shannon index (weighted by basal area) has the largest difference between models (± 0.4). The other largest variations are encountered within the models derived from composite datasets, primarily hyperspectral combined with DR LiDAR. Of these larger variations, two OLS models have the smallest R^2 values; these are the number of native trees per plot (modelled by hyperspectral and DR) and Shannon Index (weighted by basal area) (modelled by hyperspectral). For the number of tree species per plot (modelled by hyperspectral and DR), the OLS model R^2 values are higher.

The predictive power of different remote sensing inputs for compositional metrics is very similar for each metric, apart from two, (i) Simpson Index (weighted by stem count) and (ii) Simpson Index (weighted by BA). For the former, there is a disparity between model values, where hyperspectral and composite hyperspectral and FW datasets have lower values. For the latter, DR and composite hyperspectral and DR models have lower values than FW and composite hyperspectral and FW models.

In review of the seven compositional metrics two were modelled with the highest R^2 value using DR LiDAR data employing the OLS approach, and two using the hyperspectral and DR LiDAR composite using the OLS modelling approach. The three remaining metrics received high R^2 values used different remote sensing dataset combinations and modelling approaches. Hyperspectral data using an AICc approach, DR LiDAR using an OLS approach, FW LiDAR using an AICc approach, and both modelling approaches applied to the hyperspectral and FW LiDAR composite did not produce the highest R^2 values.

Table 8.20 – Forest stand composition metrics R^2 values for all datasets (all models $p < 0.05$). Blue indicates the highest R^2 value for each row.

Metric name	Hyper. OLS (R^2)	Hyper. AICc (R^2)	DR OLS (R^2)	DR AICc (R^2)	FW OLS (R^2)	FW AICc (R^2)	H+DR OLS (R^2)	H+DR AICc (R^2)	H+FW OLS (R^2)	H+FW AICc (R^2)
Number of native trees per 30x30m	0.641	0.623	0.428	0.552	0.460	0.614	0.425	0.642	0.531	0.525
Percentage of native trees	0.752	0.719	0.435	0.528	0.600	0.480	0.826	0.779	0.599	0.729
No. of tree species per 30x30m	-	0.555	0.673	0.426	0.576	0.320	0.775	0.531	0.373	0.429
Shannon Index (by stem count)	-	0.647	0.571	0.696	0.672	0.627	0.416	0.662	0.541	0.509
Simpson Index (by stem count)	0.408	0.345	-	0.589	0.546	0.488	0.527	0.587	0.371	0.371
Shannon Index (by BA)	0.188	0.620	-	0.773	0.718	0.592	0.419	0.784	0.757	0.592
Simpson Index (by BA)	-	-	0.448	0.473	0.612	0.570	0.370	0.473	0.556	0.570

Cells highlighted in blue denote the highest R^2 value for each field metric.

8.3.3 Deadwood within plot

The R^2 values for all the models produced for the prediction of the four deadwood metrics are presented in Table 8.21. Two metrics could not be modelled using OLS methods using hyperspectral and DR datasets as inputs, namely standing dead decay class and DDW volume. When considering the different statistical methods, the average difference is ± 0.135 , with a range of variation of 0.43; those AICc methods which have a lower R^2 than the OLS counterparts are illustrated in Table 8.21. The largest difference between methods (± 0.43) is found in predicting the standing deadwood decay class using composite hyperspectral and DR data. For the same metric, DR only data has a difference of ± 0.34 .

The predictive power of the different input datasets, in terms of R^2 values, is very similar for the estimation of both deadwood volume metrics; however differences are present with deadwood decay class. For standing deadwood decay class, AICc models have higher values than the equivalent models for OLS, for example DR R^2 for OLS is 0.54 whereas AICc is 0.88. This pattern is repeated for all datasets apart from the combination of hyperspectral and FW LiDAR where both OLS and AICc models have R^2 values of 0.95. For DDW decay class both hyperspectral OLS and AICc model R^2 values are lower than those derived from the other datasets. The hyperspectral model R^2 values range from 0.33-0.42, whereas the others range from 0.47-0.79.

In a review of the four deadwood metrics, two were modelled with the highest R^2 value using the hyperspectral and FW LiDAR data and OLS modelling method. The two remaining metrics received high R^2 values used different remote sensing dataset combinations and modelling approaches.

Table 8.21 – Deadwood within plot metrics R^2 values for all datasets (all models $p < 0.05$). Blue indicates the highest R^2 value for each row.

Metric name	Hyper. OLS (R^2)	Hyper. AICc (R^2)	DR OLS (R^2)	DR AICc (R^2)	FW OLS (R^2)	FW AICc (R^2)	H+DR OLS (R^2)	H+DR AICc (R^2)	H+FW OLS (R^2)	H+FW AICc (R^2)
Standing deadwood volume (m^3)	0.967	0.889	0.920	0.906	0.948	0.919	0.748	0.881	0.919	0.833
Standing deadwood decay class	-	0.903	0.541	0.883	0.47	0.589	0.476	0.902	0.955	0.955
Downed deadwood volume (m^3)	0.392	0.548	-	0.420	0.451	0.431	0.598	0.485	0.639	0.485
Downed deadwood decay class	0.335	0.418	0.786	0.472	0.749	0.569	0.681	0.580	0.750	0.717

Cells highlighted in blue denote the highest R^2 value for each field metric.

8.3.4 Regeneration and understorey metrics within the plot

The R^2 values given in Table 8.22 relate to the regression models produced for the estimation of the 12 forest understorey metrics for all five datasets. The list includes both structural and compositional understorey metrics. A number of models failed to predict some of the field metrics, three of which utilised the FW dataset using the OLS method. The DR models using the OLS method failed with two metrics. When considering the two statistical approaches, the average model variation was ± 0.17 whereas the range of variation was 0.438; AICc models which have lower R^2 values than their OLS counterparts are illustrated in Table 8.22. OLS methods generally have higher R^2 values for metrics related to seedling numbers and species.

The predictive power of different input datasets is varied throughout the regeneration and understorey subset. The most noticeable differences are found for models derived from the FW dataset, where R^2 values are lower than those of the equivalent models for five of the twelve metrics relating to the structure and composition of seedlings and vascular vegetation species, in addition to failing to predict Seedlings Shannon Index for native species. DR derived models also functioned poorly when estimating the Seedlings Shannon index for

native species. Hyperspectral derived models also performed poorly when predicting the number of bryophyte species and the percentage of bare soil within 30x30m.

In a review of the 12 understorey metrics, three were modelled with the highest R^2 value using the FW LiDAR data modelled by an AICc approach, three used the hyperspectral and FW LiDAR data composite using OLS methods and three more using AICc methods. Two metrics were modelled with the highest R^2 using hyperspectral dataset and OLS methods, two more metrics with high R^2 values were produced by DR LiDAR using an OLS approach. The two remaining metrics received high R^2 values used different remote sensing dataset combinations and modelling approaches. Only the hyperspectral data with an AICc modelling approach and hyperspectral and DR LiDAR composite with an OLS modelling approach did not generate a best fit model.

Table 8.22 – Regeneration and understorey metrics R^2 values for all datasets (all models $p < 0.05$). Blue indicates the highest R^2 value for each row.

Metric name	Hyper. OLS (R^2)	Hyper. AICc (R^2)	DR OLS (R^2)	DR AICc (R^2)	FW OLS (R^2)	FW AICc (R^2)	H+DR OLS (R^2)	H+DR AICc (R^2)	H+FW OLS (R^2)	H+FW AICc (R^2)
No. saplings per 30x30m	0.597	0.727	0.965	0.922	-	0.973	0.642	0.922	0.535	0.973
No. native saplings per 30x30m	0.646	0.678	0.890	0.896	-	0.968	0.591	0.896	0.613	0.968
Percentage of native saplings per 30x30m	0.692	0.911	0.879	0.951	0.803	0.834	0.692	0.951	0.692	0.894
No. of sapling species per 30x30m	0.663	0.798	0.905	0.938	0.702	0.954	0.494	0.922	0.848	0.895
No. seedlings per 30x30m	0.809	-	0.821	0.527	0.282	0.468	0.609	0.527	0.498	0.568
No. native seedlings per 30x30m	0.642	0.310	-	0.631	0.682	0.535	0.43	0.538	0.779	0.779
Percentage of native seedlings per 30x30m	0.816	0.856	0.920	0.585	0.683	0.432	0.763	0.577	0.605	0.865
No. of seedling species per 30x30m	0.860	0.423	0.739	0.594	0.206	0.447	0.74	0.553	0.860	0.532
Seedlings Shannon Index for native species	0.415	0.735	0.212	0.212	-	-	0.674	0.789	0.792	0.746
No. Vascular Species	0.855	0.437	-	0.746	0.194	0.194	-	0.547	0.404	-
No. Bryophyte Species	0.439	0.439	0.835	0.858	0.782	0.457	0.737	0.625	0.687	0.687
Percentage of bare soil within 30x30m	0.518	0.518	0.761	0.737	0.862	0.749	0.761	0.793	0.840	0.749

Cells highlighted in blue denote the highest R^2 value for each field metric.

8.3.5 Metrics required for condition index construction

The two metrics integral to index construction and their associated R^2 values are given in Table 8.23. Given the structural nature of these metrics and time restraints, only LiDAR and composite datasets were used in the regression analyses, and only AICc for composite models. Of the two analyses which used OLS and AICc, R^2 values were again similar, with AICc values being lower than those of OLS in 1 out of 8 models. Sum of DBH differences received high R^2 values overall, whereas the index of vertical separation was more variable. The index of vertical separation was higher for AICc methods, and was highest for DR LiDAR and both composite (i.e. hyperspectral & DR and hyperspectral & FW) derived models.

Table 8.23 – Index Inputs metrics R^2 values for all datasets (all models $p < 0.05$). Blue indicates the highest R^2 value for each row.

Metric names	DR OLS (R^2)	DR AICc (R^2)	FW OLS (R^2)	FW AICc (R^2)	H+DR OLS (R^2)	H+DR AICc (R^2)	H+FW OLS (R^2)	H+FW AICc (R^2)
Sum of the DBH differences between neighbouring trees (cm)	0.873	0.924	0.912	0.714	-	0.950	-	0.864
Index of Vertical Separation	0.526	0.743	0.467	0.601	-	0.730	-	0.736

Cells highlighted in blue denote the highest R^2 value for each field metric.

8.4A review and summary of model prediction accuracy of all remote sensing datasets

This section contains the results of the regression models derived through remote sensing means as compared with the 2012 field data measurements. RMSE and NRMSE were calculated for each of the remote sensing derived models to identify which of the ten models was the ‘best’ the predictor of the field data. Regression derived models were chosen with the highest R^2 values ($p < 0.05$) and the lowest RMSE and NRMSE values. An overview of which of the ten models was the best predictor for each field metric is presented in Table 8.24 for overstorey metrics and in Table 8.25 for the understorey, for the groupings of: (i) DR LiDAR, hyperspectral, and the composite DR and hyperspectral dataset metrics, and (ii) FW LiDAR, hyperspectral and composite FW and hyperspectral dataset metrics. Two field-level metrics could not be adequately modelled; these were the percentage of native saplings within the plot and the number of bryophyte species.

Tables 8.26 and 8.27, show that the RMSE and NRMSE values are similar for the majority of cases between the two main groups: (i) DR LiDAR and hyperspectral metrics and (ii) FW LiDAR and hyperspectral metrics. In comparison, NRMSE values typically vary by ± 0.04 of one another for each metric. The majority of the NRMSE values are below 0.4. Five metrics have higher values for both model sets, these are:

- (1) Percentage very big trees (DBH >80cm);
- (2) Shannon Index (by basal area);
- (3) Percentage of native seedlings per 30x30m;
- (4) Seedlings Shannon Index for native species;
- (5) Percentage of bare soil within 30x30m; and

The percentage very big trees (DBH >80cm) estimate produced an NRMSE above 1 however, i.e. beyond the range of the field data. The Shannon Index (by basal area) estimates produced NRMSE values of greater than 0.5. For both these metrics the error was considered too high to be used. It should also be noted that the Simpson index (by stem count) was best predicted by applying the index calculation directly to the hyperspectral data.

The choices of best statistical models for groups (i) DR LiDAR, hyperspectral, and composite DR and hyperspectral dataset metrics, and (ii) all FW LiDAR, hyperspectral, and composite FW and hyperspectral dataset metrics, in addition to how many of each were identified as the best model overall, are presented in Table 8.26. Overall, LiDAR and composite datasets provided the most predicative ability. Out of the remaining 35 metrics this is best summarised as:

- 9 DR LiDAR only;
- 15 FW LiDAR only;
- 2 hyperspectral only;
- 4 combined DR LiDAR and hyperspectral;
- 5 combined FW LiDAR and hyperspectral.

In particular, FW LiDAR provided the most input to predicating 20 (15 LiDAR and 5 composite) (57%) of the field-level metrics, in comparison to 13 for DR (9 LiDAR and 4 composite) (37%), and 2 for hyperspectral models (6%).

A summary of which field-level metrics were best predicted by which dataset and regression model is available in Table 8.27. The model equations and coefficients are given in Table 8.28. Both LiDAR datasets predict many of the forest structural variables. There are also five overstorey and understorey compositional metrics best described without any input from hyperspectral data, for example the number of tree species and number of vascular species. Hyperspectral data on their own best estimates two measures of species diversity and an understorey metric, downed deadwood decay class. The combination of DR or FW LiDAR with hyperspectral datasets are best predictors of some overstorey and understorey compositional metrics, for example sapling and seedling species numbers.

Table 8.24 – The best regression models identified through interrogation of RMSE and NRMSE values. Part 1: overstorey composition. [DR = DR LiDAR; FW = FW LiDAR; Hyp = Hyperspectral dataset metrics]

Field metric	(1)DR, (2)Hyp, and (3)DR& Hyp Model	(1)DR, (2)Hyp, and (3)DR& Hyp RMSE	(1)DR, (2)Hyp, and (3)DR& Hyp NRMSE	(1)FW, (2)Hyp, and (3)FW& Hyp Model	(1)FW, (2)Hyp, and (3)FW& Hyp RMSE	(1)FW, (2)Hyp, and (3)FW& Hyp NRMSE	Best model overall
Number of trees per 30x30m	DR AICc	18.90	0.256	FW OLS	15.974	0.216	FW OLS
Number of native trees per 30x30m	DR AICc	15.04	0.289	FW and Hyper. AICc	12.64	0.243	FW and Hyper. AICc
Percentage of native trees	DR AICc	28.75	0.288	FW AICc	29.74	0.297	DR AICc
No. of tree species per 30x30m	DR AICc	1.50	0.250	FW AICc	1.66	0.276	DR AICc
Average tree spacing (m)	DR OLS	1.71	0.330	FW OLS	1.32	0.398	DR OLS
STDev. of tree spacing (m)	DR OLS	0.58	0.695	FW OLS	0.39	0.475	FW OLS
Mean DBH (cm)	DR OLS	8.38	0.287	FW and Hyper. AICc	12.63	0.433	DR OLS
STDev. of DBH (cm)	DR OLS	10.80	0.217	FW AICc	9.10	0.187	FW AICc
Basal area within 30x30m	DR AICc	0.83	0.257	FW OLS	0.91	0.283	DR AICc
Percentage big trees(DBH 40>x<80cm)	DR and Hyper. OLS	20.03	0.286	FW OLS	19.93	0.285	FW OLS
Percentage very big trees (DBH >80cm)	DR AICc	7.04	1.162	FW OLS	10.31	1.701	DR AICc
Percentage big trees (DBH >40cm)	DR and Hyper. OLS	26.22	0.375	FW AICc	21.89	0.313	FW AICc
Shannon index (by stem count)	DR and Hyper. OLS	0.44	0.308	FW and Hyper OLS	0.52	0.365	DR and Hyper. OLS
Simpson index (by stem count)	Hyper. (Direct)	0.28	0.217	FW AICc	0.29	0.221	Hyper (Direct)
Shannon index (by BA)	Hyper. OLS	0.45	0.589	Hyper. OLS	0.45	0.589	Hyper. OLS
Simpson index (by BA)	DR and Hyper. OLS	0.62	0.217	FW AICc	0.79	0.275	DR and Hyper. OLS
Average height of the first live branch (m)	DR and Hyper. AICc	2.66	0.245	FW OLS	2.59	0.239	FW OLS
STDev. of the height of the first live branch (m)	DR AICc	1.66	0.322	FW and Hyper. AICc	1.87	0.363	DR AICc
Average crown horizontal area (m ²)	DR AICc	23.53	0.270	FW OLS	24.39	0.280	DR AICc
STDev. crown horizontal area (m ²)	DR and Hyper. AICc	17.77	0.311	FW AICc	17.59	0.308	FW AICc
Total crown horizontal area (m ²)	DR and Hyper. OLS	510.52	0.354	Hyper. AICc	538.76	0.373	DR and Hyper. OLS

Table 8.25 – The best regression models identified through interrogation of RMSE and NRMSE values. Part 2: understorey composition. [DR = DR LiDAR; FW = FW LiDAR; Hyp = Hyperspectral dataset metrics]

Field metric	(1)DR, (2)Hyp, and (3)DR& Hyp Model	(1)DR, (2)Hyp, and (3)DR& HypRM SE	(1)DR, (2)Hyp, and (3)DR& HypNR MSE	(1)FW, (2)Hyp, and (3)FW& Hyp Model	(1)FW, (2)Hyp, and (3)FW& HypRM SE	(1)FW, (2)Hyp, and (3)FW& HypNR MSE	Best model overall
<i>Standing deadwood volume (m³)</i>	<i>DR OLS</i>	1.50	0.181	<i>FW AICc</i>	1.36	0.165	<i>FW AICc</i>
<i>Standing deadwood decay class</i>	<i>DR and Hyper. AICc</i>	0.22	0.325	<i>FW AICc</i>	0.19	0.290	<i>FW AICc</i>
<i>Downed deadwood volume (m³)</i>	<i>DR OLS</i>	2.74	0.300	<i>FW OLS</i>	2.49	0.270	<i>FW OLS</i>
<i>Downed deadwood decay class</i>	<i>Hyper. OLS</i>	0.16	0.339	<i>Hyper. OLS</i>	0.16	0.339	<i>Hyper. OLS</i>
<i>No. saplings per 30x30m</i>	<i>DR AICc</i>	24.41	0.230	<i>FW and Hyper. OLS</i>	24.42	0.230	<i>DR AICc</i>
<i>No. native saplings per 30x30m</i>	<i>DR and Hyper. OLS</i>	5.73	0.151	<i>FW and Hyper. OLS</i>	9.17	0.24	<i>DR and Hyper. OLS</i>
<i>Percentage of native saplings per 30x30m</i>	-	-	-	-	-	-	-
<i>No. of sapling species per 30x30m</i>	<i>Hyper. OLS</i>	1.12	0.373	<i>FW and Hyper. OLS</i>	1.02	0.342	<i>FW and Hyper. OLS</i>
<i>No. seedlings per 30x30m</i>	<i>DR OLS</i>	295.24	0.315	<i>FW and Hyper. OLS</i>	260.46	0.278	<i>FW and Hyper. OLS</i>
<i>No. native seedlings per 30x30m</i>	<i>DR OLS</i>	239.08	0.332	<i>FW OLS</i>	247.84	0.340	<i>DR OLS</i>
<i>Percentage of native seedlings per 30x30m</i>	<i>DR and Hyper. OLS</i>	42.17	0.422	<i>FW OLS</i>	39.72	0.397	<i>FW OLS</i>
<i>No. of seedling species per 30x30m</i>	<i>DR OLS</i>	3.16	0.243	<i>FW and Hyper. AICc</i>	3.09	0.238	<i>FW and Hyper. AICc</i>
<i>Seedlings Shannon index for native species</i>	<i>DR AICc</i>	0.73	0.450	<i>FW and Hyper. AICc</i>	0.70	0.436	<i>FW and Hyper. AICc</i>
<i>No. vascular species</i>	<i>DR and Hyper. AICc</i>	1.99	0.331	<i>FW OLS</i>	1.70	0.283	<i>FW OLS</i>
<i>No. bryophyte species</i>	-	-	-	-	-	-	-
<i>Percentage of bare soil within 30x30m</i>	<i>DR AICc</i>	44.04	0.464	<i>FW OLS</i>	39.82	0.419	<i>FW OLS</i>
<i>Sum of the DBH differences between neighbouring trees (cm)</i>	<i>Hyper. AICc</i>	23.54	0.569	<i>FW AICc</i>	19.54	0.389	<i>FW AICc</i>
<i>Index of vertical separation</i>	<i>DR OLS</i>	0.16	0.323	<i>FW OLS</i>	0.13	0.262	<i>FW OLS</i>

Table 8.26 – Summary of best statistical model choices [DR = DR LiDAR; FW = FW LiDAR; Hyp = Hyperspectral dataset metrics]

Dataset	(1)DR, (2)Hyp, and (3)DR&Hyp models count	(1)FW, (2)Hyp, and (3)FW&Hyp models count	Best model overall count
DR OLS	11	-	3
DR AICc	13	-	6
FW OLS	-	14	9
FW AICc	-	11	6
Hyperspectral OLS	4	3	1
Hyperspectral AICc	1	2	0
DR and Hyper. OLS	5	-	4
DR and Hyper. AICc	3	-	0
FW and Hyper. OLS	-	2	2
FW and Hyper. AICc	-	4	3
Hyperspectral Direct	1	1	1

Table 8.27– A breakdown of the regression models which best predicted each field-level metric

Dataset	Best model overall count	Field-level metrics names
DR OLS	3	Average tree spacing; Mean DBH (cm); No. native seedlings per 30x30m;
DR AICc	6	Percentage of native trees per 30x30m; No. of tree species per 30x30m; Basal area within 30x30m (m ²); STDev of the height of the first live branch (m); Average crown horizontal area (m ²); No. saplings per 30x30m;
FW OLS	9	Number of trees per 30x30m; STDev. of tree spacing; Percentage big trees (DBH 40>x<80cm); Average height of the first live branch (m); Downed deadwood volume (m ³); Percentage of native seedlings per 30x30m; No. of vascular species; Percentage of bare soil; Index of vertical separation.
FW AICc	6	STDev. DBH (cm); Percentage of big trees (DBH>40cm); STDev. of crown horizontal area (m ²); Standing deadwood volume (m ³); Standing deadwood decay class; Sum of the DBH differences between neighbouring trees (cm);
Hyper. OLS	1	Downed deadwood class;
Hyper. AICc	0	-
DR and Hyper. OLS	4	Shannon index (by stem count); Simpson index (by BA); Total crown horizontal area (m ²); No. native saplings per 30x30m;
DR and Hyper. AICc	0	-
FW and Hyper. OLS	2	No. sapling species per 30x30m No. seedling per 30x30m;
FW and Hyper. AICc	3	Number of native trees per 30x30m; No. seedling species per 30x30m; Seedling Shannon index for native species;
Hyperspectral Direct	1	Simpson index (by stem count);

Table 8.28 – Regression model list for a combination of the ‘best’ FW, DR, hyperspectral or composite models. The table displays the regression equation coefficients (B_0 , $B_1...B_n$) and predictor metrics (X_1 , $X_2...X_n$) for each derived model. The two red lines indicates models which could not be adequately modelled.

Y	B0	B1	X1	B2	X2	B3	X3	B4	X4	B5	X5
Number of trees per 30x30m	40.45	-0.33	TIF_m_area	2.12	ht_p20_2	-0.45	kurngamp	3.01	skwnght	-	-
Number of native trees per 30x30m	47.92	1.45	sr12_Min	-1.24	wd_p45	-	-	-	-	-	-
Percentage of native trees	125.52	-2.68	TIF_TreeNo	-	-	-	-	-	-	-	-
No. of tree species per 30x30m	6.78	-0.14	p050	-0.0008	SIN_Var_2	-	-	-	-	-	-
Average tree spacing (m)	-2.23	0.58	TIF_Space	0.61	amp_p20	-	-	-	-	-	-
STDev. Of tree spacing (m)	1.32	0.01	TIF_m_area	-0.09	meanht_2	0.03	kurwid	-	-	-	-
Mean DBH (cm)	16.86	2.69	TIF_space	0.004	TIF_meanCV	-	-	-	-	-	-
STDev. of DBH (cm)	14.14	-1.05	amp_p65	0.19	TIF_m_area	0.95	wd_p25_2	-	-	-	-
Basal area within 30x30m	3.21	0.07	p050_2	-0.15	TIF_Space	-	-	-	-	-	-
Percentage big trees (DBH 40>x<80cm)	-40.72	-6.25	kurgamp_2	4.03	mednht	3.72	TIF_Space	-	-	-	-
Percentage very big trees (DBH >80cm)	15.7	-0.53	TIF_TreeNo	-	-	-	-	-	-	-	-
Percentage big trees (DBH >40cm)	-54.95	10.22	adevht_2	5.69	TIF_Space	-	-	-	-	-	-
Shannon Index (by stem count)	1.56	0.4	rgri1_Min	0.67	rgri2_Min	0.3	SHN_Ske	-0.92	SIA_Ske	-	-
Simpson Index (by BA)	0.88	12.83	ndvi2_STD	0.01	no_nat_t	-0.3	SHA_Ske_2	0.008	SIA_Med	0.06	sr12_Min
Shannon Index (by BA)	0.3	0.02	no_nat_t	-	-	-0.3	SHA_Ske_2	0.008	SIA_Med	-	-
Simpson Index (by BA)	-0.88	12.83	ndvi2_STD	0.01	no_nat_t	-0.3	SHA_Ske_2	0.008	SIA_Med	0.06	sr12_Min
Average height of the first live branch (m)	7.48	0.76	loft_max_gpp	-0.52	TIF_Space	-0.00006	TIF_totalCV	-0.004	varngamp_2	-	-
STDev. of the height of the first live branch (m)	4.95	0.16	SHA_Max_2	-0.64	TIF_STD	-0.006	TIF_to_area	-	-	-	-
Average crown horizontal area (m2)	-44.39	1.76	SIN_Kur	6.31	TIF_Space	0.06	TIF_to_area	-	-	-	-
STDev. Crown horizontal area (m2)	136.78	-3.67	amp_p85	-	-	-	-	-	-	-	-
Total Crown horizontal area (m2)	1064.44	0.78	ari2_Sum	-0.29	SIA_Var_2	-	-	-	-	-	-
Standing dead wood volume (m3)	-8.63	1.4	skwngamp	2.37	TIF_STD	0.35	wd_p25_2	-	-	-	-
Standing dead wood decay class	1.22	-0.06	amp_p35_2	-0.016	ht_p20_2	-	-	-	-	-	-
Downed dead wood volume (m3)	-13.85	3.17	skwngamp	1.77	adevht_2	-	-	-	-	-	-
Downed dead wood decay class	0.53	0.01	sr12_Mean	-	-	-	-	-	-	-	-
No. saplings per 30x30m	27.73	-2.79	p060_2	2.68	p060_2	-0.42	SIA_Mean	-	-	-	-
No. native saplings per 30x30m	-15.63	1.73	mtresr12_Max	14.21	SIG_Ske	-	-	-	-	-	-
Percentage of native saplings per 30x30m	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL
No. of sapling species per 30x30m	1.64	-0.34	ari2_Min	2.08	ndvi2_Min	-0.1	no_nat_t	-	-	-	-
No. seedlings per 30x30m	-248	8.58	varnght	59.34	spec_no	-	-	-	-	-	-
No. native seedlings per 30x30m	-89.24	74.71	TIF_STD	0.05	TIF_meanCV	-	-	-	-	-	-
Percentage of native seedlings per 30x30m	139.09	-1.7	medgamp_2	-1.23	ht_p25_2	-	-	-	-	-	-
No. of seedling species per 30x30m	-4.99	0.09	sr12_Max	0.24	stdgwid	-	-	-	-	-	-
Seedlings Shannon Index for native species	-1.09	0.13	adevht_2	0.15	spec_no	-	-	-	-	-	-
No. Vascular Species	2.67	0.03	varnght_2	-	-	-	-	-	-	-	-
No. Bryophyte Species	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL
Percentage of bare soil within 30x30m	-23.23	3.66	ht_p20_2	7.67	loft_can_depth	-16.51	TIF_STD	-	-	-	-
Sum of the DBH differences between neighbouring trees (cm)	-34.51	0.59	adevwd_loft	1.26	amp_p90_loft	0.86	ht_p20	-3.4	kurngwid	-	-
Index of Vertical Separation	0.966	0.07	skwnght_2	-0.005	amp_p85	-	-	-	-	-	-

8.5 Summary of findings

In summary, this chapter began by presenting the results of the hyperspectral combined with DR LiDAR, and hyperspectral combined with FW LiDAR field metric predictive models. For the former, models could be produced for all 39 indicator metrics using both OLS and AICc methods, apart from two using the OLS method (the sum of diameter differences and the index of vertical separation metrics). Thus, OLS performed better for 12 out of 37 models in terms of R^2 value. Eight out a total of 37 indicators had NRMSE values above 0.5 for both models, these were: average tree NN spacing; standard deviation of tree NN spacing; mean DBH; Percentage very big trees; Shannon Index (by basal area); Percentage of native saplings; No. bryophyte Species; and Index of vertical separation. For OLS, the standard deviation of the height of the first live branch and DDW volume had NRSME values above 0.5. For AICc the standing deadwood volume had an NRMSE value above 0.5. Overall OLS performed better for 12 indicators out of the remaining 31 metrics.

Models for hyperspectral combined with FW LiDAR could be produced for 37 indicator metrics for OLS and 39 for AICc methods. OLS methods could not estimate the sum of diameter differences and the index of vertical separation metrics, and AICc methods could not estimate the number of vascular species. OLS performed better for 14 out of 37 models in terms of R^2 . Five out a total of 39 indicators had NRMSE values above 0.5 for both models, these were: average tree NN spacing; standard deviation tree NN spacing; Percentage very big trees; Shannon Index (by BA); and Percentage of native saplings. For OLS mean DBH, DDW volume, and DDW decay class NRMSE values were above 0.5, whilst for AICc basal area, standard deviation of crown horizontal area and standing deadwood volume NRMSE values were above 0.5. Overall OLS performed better for 13 out of the remaining 34 metrics.

Comparison of the models derived from all the remote sensing datasets for each of the indicators showed that the predictive power of different remote sensing inputs for structural metrics varied, and most notably between hyperspectral and LiDAR datasets in terms of R^2 value. Comparisons between predicted metrics and validation field plot data indicated automated ITC detection performed very similarly between TIFFS and eCongition solutions. The best estimates of mean canopy height came from DR LiDAR average TIFFS ITC data, with FW leaf-on LiDAR dominant height giving very similar results. Canopy layer estimates were also very similar between DR and FW LiDAR datasets. Canopy openness was best estimated by FW leaf-on LiDAR data.

The best regression models for the field indicator metrics in relation to RMSE and NRMSE were presented for all relevant combinations of hyperspectral and DR, and hyperspectral and FW LiDAR inputs. The best models from the two main groups were often very similar. The percentage of very big trees (DBH>80cm), Shannon index (by basal area), percentage of native saplings and the number of bryophyte species could not be adequately modelled. The subset of the best models, out of a total of 35, contained (i) 20 models from the FW and FW combined with hyperspectral datasets; (ii) 13 models from the DR LiDAR and DR combined with hyperspectral datasets; and (iii) 2 models from hyperspectral datasets only. These models will subsequently be used to estimate forest condition over the whole study site.

Chapter 9 – Mapping forest metrics and condition indices

This chapter presents maps of the 37 forest structure and compositional metrics derived at a 30x30m resolution from the best regression model results identified in Chapter 8. This also includes the two metrics which received poor NRMSE values (i.e. >0.4), the percentage very big trees (DBH $>80\text{cm}$) and the Shannon index (by basal area). Condition indices were calculated from these remote sensing derived 30x30m field-level metrics and were subsequently validated with field-measurement indices. These various indices incorporate both structural and compositional metrics in various numbers and combinations, relating to horizontal and vertical distributions within the fieldwork plots. As with the estimates of field metrics, the condition indices derived through remote sensing data were mapped across the whole study site.

9.1 Mapping of forest metrics

As stated in section 4.7, the regression models could be applied for the entirety of the study site, provided there was coincident airborne remote sensing data coverage to provide input metrics. The following section documents the mapping outputs for the best models for all 37 field metrics derived from the remote sensing datasets (i) hyperspectral, (ii) DR LiDAR, (iii) FW LiDAR, (iv) composite hyperspectral and DR LiDAR, or (v) composite hyperspectral and FW LiDAR, as identified in chapter 8.4, for a subset of the total study area in order to illustrate features at the plot and stand scales centred upon the Frame Wood area, with elements of Tantany Woods, Frame-Heath and Hawkhill inclosures visible. This area contains a mixture of semi-natural ancient woodland, such as those located in Frame and Tantany Woods, and managed plantation woodland, such as that in the Frame-Heath and Hawkhill inclosures. Figure 9.1 illustrates the subset area comprising a 1x1m resolution CHM, and the primary species as identified by FC inventory data. Table 9.1 reports the map name, figure number and description of each of the 37 predicted metrics.

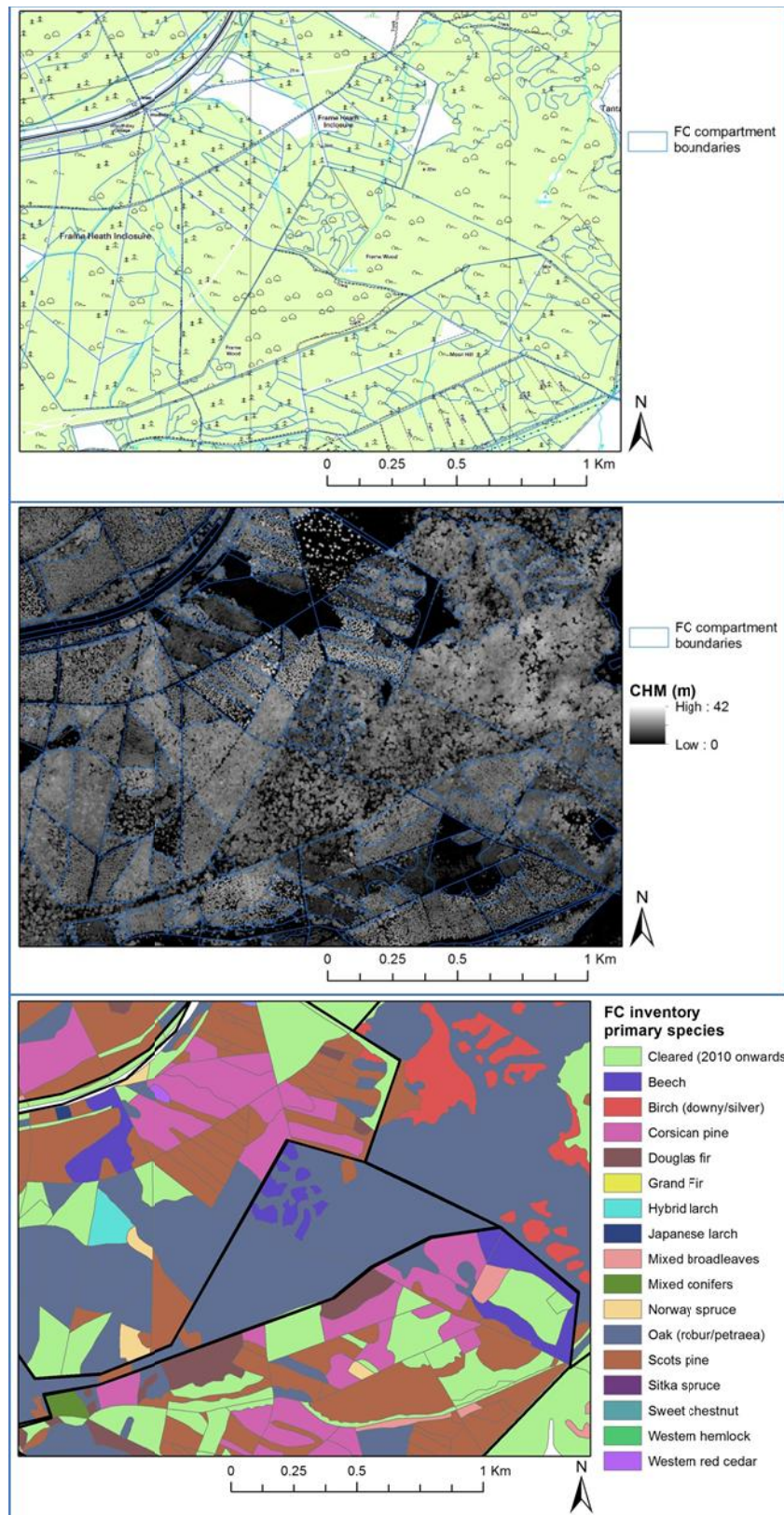


Figure 9.1 – The subset area for presenting predicted field-level metrics over wide areas. Frame Wood and Tantany Wood in addition to Frame-Heath and Hawkhill inclosures intersect this area. Ordnance survey mapping, 1x1m raster nCHM, and FC inventory primary species classification are presented for reference. Base Map layer (top) is © Crown Copyright/database right 2010. An Ordnance Survey/EDINA supplied service.

It should be noted that there was an obvious error present in each of the maps produced for this section which could not be rectified. This error lies to the east of the subset area and runs north to south. This error was caused by errors in mosaicing the extracted FW LiDAR metrics together caused by duplicate grid cells in the shapefiles. Given more time this issue could be rectified, however this minor error does not detract from the overall assessment of the mapping approaches.

A comparison of 10 sample metrics estimated from (i) DR, hyperspectral or DR and hyperspectral composite; or (ii) FW, hyperspectral or FW and hyperspectral composite models is available in Appendix F section F.1.

Table 9.1 – Summary of mapping results

Metric name	Description
<i>Number of trees per 30x30m (Figure 9.2)</i>	The map indicates higher stem densities within the plantation inclosures. A relatively even tree stem density can be observed through much of the semi-ancient Frame woodland. Variability occurs within Tantany wood where there are pockets of birch in amongst the beech and oak dominated areas.
<i>Number of native trees per 30x30m (Figure 9.3)</i>	The pattern of higher numbers is preserved within the semi-ancient woodland areas, the differences are evident however in plantation woodlands, due to the higher proportions of non-native tree species.
<i>Percentage of native trees per 30x30m (Figure 9.4)</i>	Estimates of the total percentage of trees of a native species per 30x30m area broadly identify areas of coniferous plantation and deciduous semi-ancient woodland.
<i>No. of tree species per 30x30m (Figure 9.5)</i>	A maximum of seven tree species were identified within the 30x30m gridded area. Semi-ancient woodlands exhibit relatively few species. Plantation areas can exhibit only one, or a maximum of seven, but these areas are noticeably stratified by compartment boundaries.
<i>Average tree NN spacing (m) (Figure 9.6)</i>	Average tree spacing between nearest neighbours is variable across the semi-ancient woodlands. The coniferous plantation areas are noticeably stratified by compartment boundaries, where different compartments may exhibit different average spacing, but are generally uniform within.
<i>STDev. of tree NN spacing (m) (Figure 9.7)</i>	Within semi-ancient deciduous woodland the standard deviation of nearest neighbour tree spacing is variable, even between the two semi-ancient woodlands, Tantany and Frame. The plantation woodland areas do not exhibit a lot of variability
<i>Mean DBH (cm) (Figure 9.8)</i>	Generally higher, but more variable mean DBH values are found in the semi-ancient deciduous woodland. DBH variation is limited in the coniferous woodlands and is noticeably confined by compartment boundaries.
<i>STDev. of DBH (cm) (Figure 9.9)</i>	Generally higher and more variable standard deviation DBH values are found in the semi-ancient deciduous woodland. DBH variation is limited in the coniferous woodlands and is noticeably confined by compartment boundaries.
<i>Basal area within 30x30m (Figure 9.10)</i>	Basal area is significantly higher in coniferous plantation woodland, and is typically uniform within each compartment area. Lower and more variable basal area values are present within semi-ancient woodlands.

Table 9.1 (continued)

Metric name	Description
<i>Percentage big trees (DBH 40>x<80cm) (Figure 9.11)</i>	The percentage of big trees is values are typically higher in certain plantation areas and parts of the semi-ancient woodlands. Many plantation compartments contain very low values (~0%).
<i>Percentage very big trees (DBH >80cm) (Figure 9.12)</i>	Much higher values are detected in deciduous woodland and none or very low values are detected in coniferous woodlands. Note RMSE value for this model high.
<i>Percentage big trees (DBH >40cm) (Figure 9.13)</i>	The percentage of big trees is typically higher in certain plantation areas and parts of the semi-ancient woodlands. Many plantation compartments contain very low values (~0%).
<i>Shannon index (by stem count) (Figure 9.14)</i>	The highest SH values are found in Tantany Wood, with more variable values found in Frame Wood. Some coniferous dominated compartments have high values, but the majority have low SH values.
<i>Simpson index (by stem count) (Figure 9.15)</i>	The highest SI values are found in the deciduous woodland. Some coniferous dominated compartments have high values, but the majority have low SI values.
<i>Shannon index (by BA) (Figure 9.16)</i>	Higher values are located within plantation woodlands, i.e. due to higher basal area values.
<i>Simpson index (by BA) (Figure 9.17)</i>	Higher values are located within plantation compartments and certain deciduous woodlands, i.e. due to higher basal area values. Although there are a number of coniferous compartments with very low values.
<i>Average height of the first live branch (m) (Figure 9.18)</i>	The average height is consistently higher in coniferous areas, and lower in deciduous areas.
<i>STDev. of the height of the first live branch (m) (Figure 9.19)</i>	The standard deviation of the height of the first live branch is highly variable across the whole site; however larger values are encountered in coniferous compartment areas.
<i>Average crown horizontal area (m²) (Figure 9.20)</i>	Semi-ancient deciduous woodlands contain much larger mean crown areas and variation in values. Conifer areas contain lower and less variable values.
<i>STDev. crown horizontal area (m²) (Figure 9.21)</i>	Semi-ancient deciduous woodlands contain much larger standard deviation of crown area and variation in values. Conifer areas contain lower and less variable values.
<i>Total crown horizontal area (m²) (Figure 9.22)</i>	The total crown horizontal area shows a different pattern to average and standard deviation of crown area. Larger values are generally found in deciduous woodland areas, Frame Wood however has lower values, potentially related to the larger canopy gaps evident in the CHM.
<i>Standing deadwood volume (m³) (Figure 9.23)</i>	Standing deadwood volume is higher in deciduous woodlands, and also exhibits more variability between cells. Low deadwood volume is located within the coniferous plantation areas.
<i>Standing deadwood decay class (Figure 9.24)</i>	Deadwood decay class is highly variable across the woodlands, but obvious differences between woodland compartments are evident. Generally decay class is higher in many deciduous areas, apart from Tantany Wood, and lower in coniferous compartments.
<i>Downed deadwood volume (m³) (Figure 9.25)</i>	Downed deadwood volume follows a similar pattern to standing deadwood volume, where deciduous areas have higher values. Although TantanyWood exhibits a change from east to west, from high to low. Deadwood volume in coniferous areas is generally low, but varies between compartments.

Table 9.1 (continued)

Metric name	Description
<i>Downed deadwood decay class (Figure 9.26)</i>	Overall deadwood decay class is highest in deciduous areas, which is significantly higher than coniferous areas. The various coniferous compartments generally differ in terms of decay class.
<i>No. saplings per 30x30m (Figure 9.27)</i>	Sapling number is generally low in field work, which is repeated in the estimation from remote sensing data. Larger counts are located in specific compartments of both coniferous and deciduous woodlands.
<i>No. native saplings per 30x30m (Figure 9.28)</i>	This map follows a similar pattern to the total number of saplings with larger values and greater variability in deciduous woodlands. Coniferous compartments are generally more uniform, and can also vary in count value.
<i>No. of sapling species per 30x30m (Figure 9.29)</i>	Semi-ancient woodlands exhibit relatively few sapling species. Plantation areas can exhibit general lower numbers, but these areas are noticeably stratified by compartment boundaries.
<i>No. seedlings per 30x30m (Figure 9.30)</i>	Seedlings count is variable across the whole site but is generally stratified by woodland compartment.
<i>No. native seedlings per 30x30m (Figure 9.31)</i>	Native seedlings count is variable across the whole site but is generally stratified by woodland compartment.
<i>No. of seedling species per 30x30m (Figure 9.32)</i>	Generally larger values are found within deciduous areas. In addition conifer areas are generally stratified by woodland compartment.
<i>Seedlings Shannon index for native species (Figure 9.33)</i>	SH index values are generally higher in coniferous areas and very low in deciduous areas.
<i>No. vascular species (Figure 9.34)</i>	The number of vascular plants is generally higher in deciduous areas. Coniferous areas can be variable from low to high.
<i>Percentage of bare soil within 30x30m (Figure 9.35)</i>	Higher amounts of bare ground/soil are found in deciduous areas, however a number of conifer compartments have high values also.
<i>Sum of the DBH differences between neighbouring trees (cm) (Figure 9.36)</i>	The differences in DBH value are higher in the plantation woodlands, and a limited number of deciduous areas. Deciduous areas are generally less different.
<i>Index of vertical separation (Figure 9.37)</i>	This metric is related to the variance of the largest and smallest DBH values within the 30x30m cell. This is found to be highest generally in deciduous areas, especially Frame Wood, where oak and birch woodlands intermingle. Coniferous woodland index values are typically low.

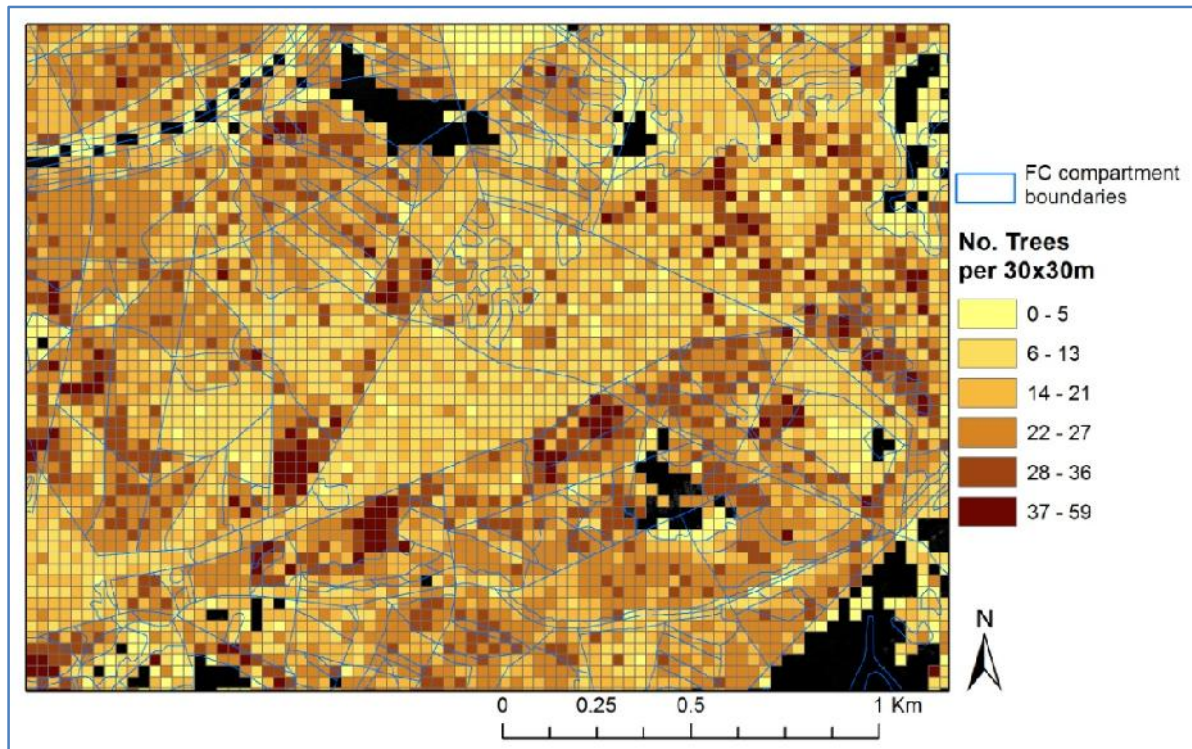


Figure 9.2 – The predicted number of tree stems per 30x30m

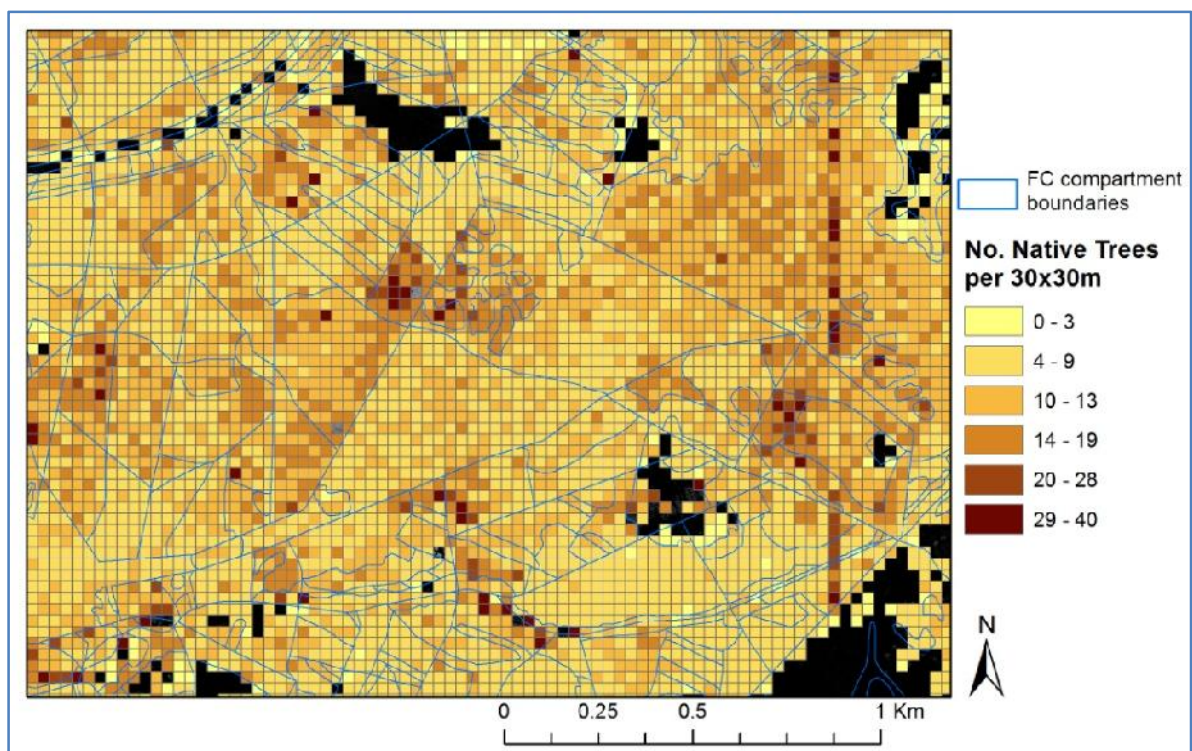


Figure 9.3 – The predicted number of tree stems of native species per 30x30m

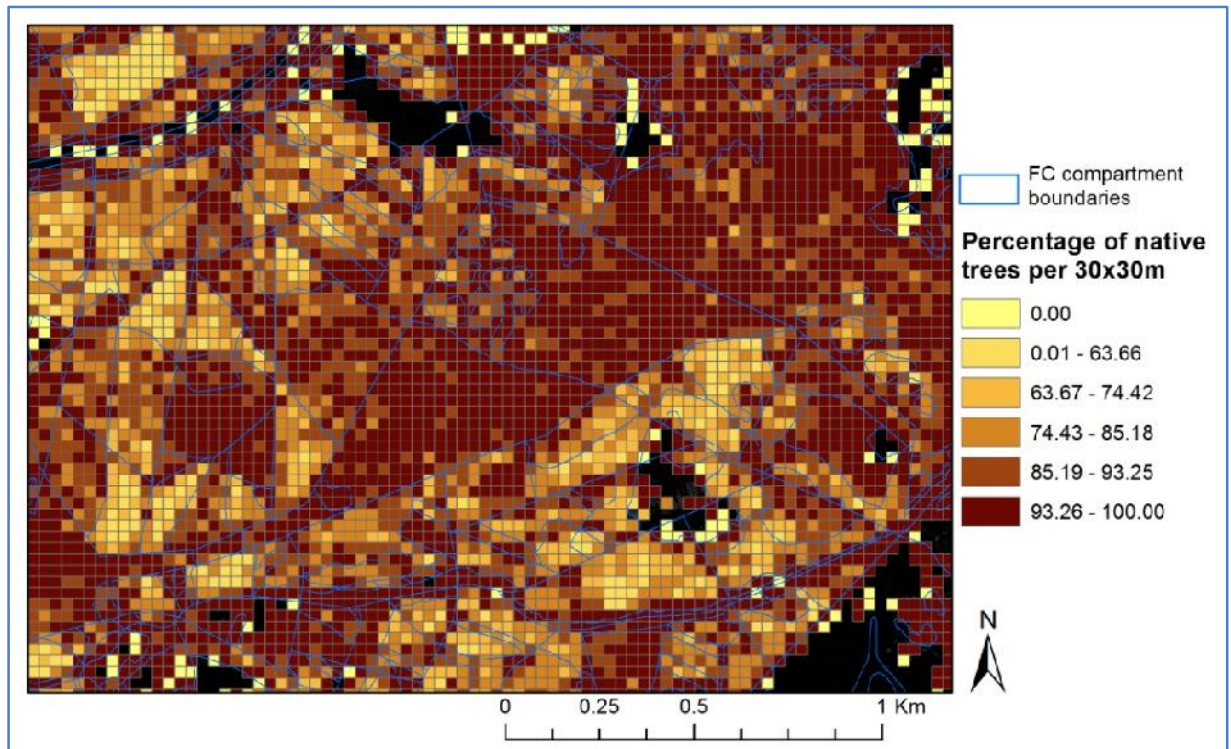


Figure 9.4 – The predicted percentage of tree stems of native species per 30x30m

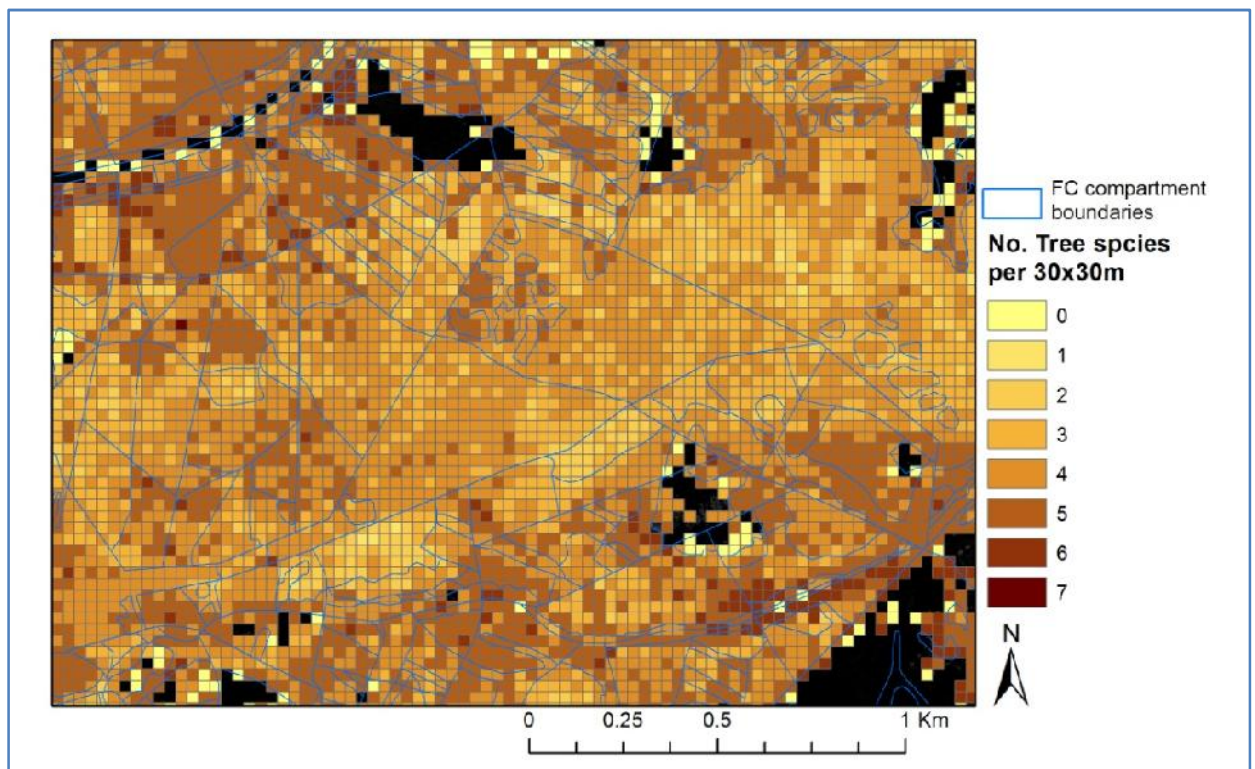


Figure 9.5 – The predicted number of tree species per 30x30m

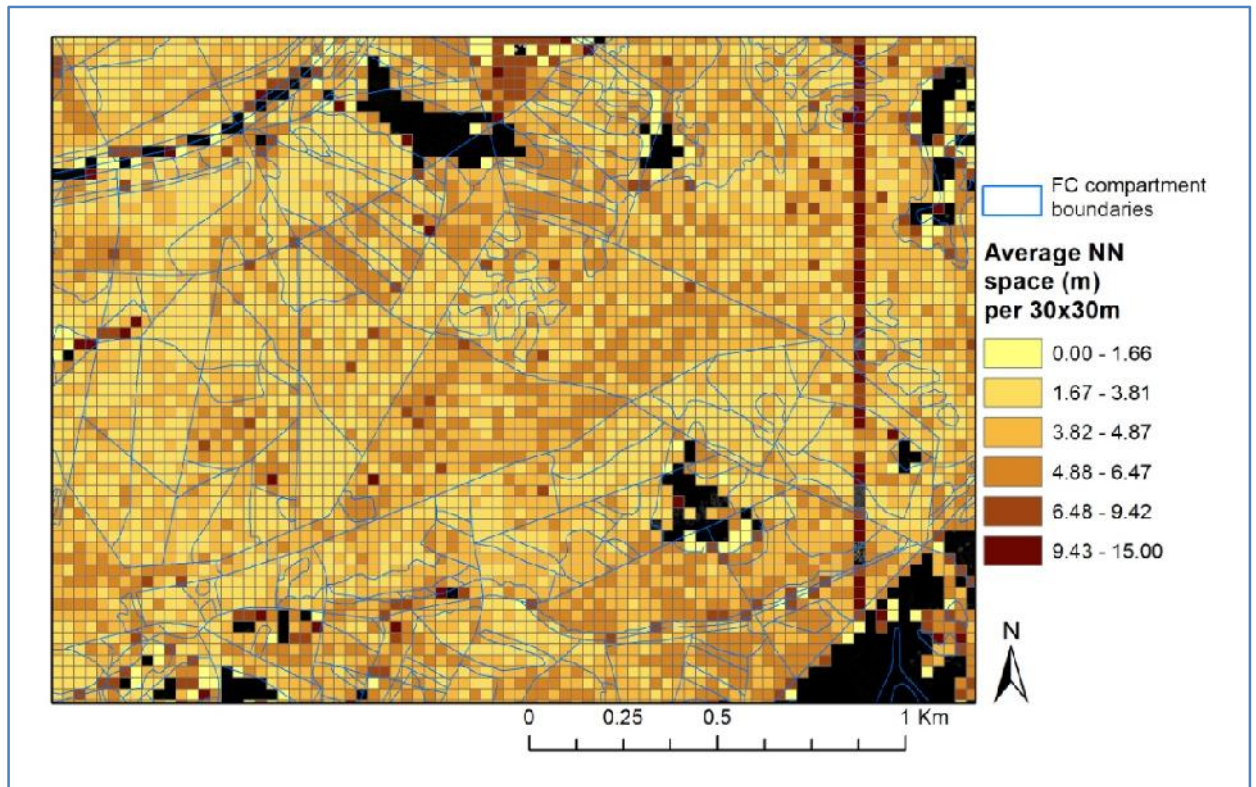


Figure 9.6 – The predicted average of tree stem nearest neighbour spacing (m) per 30x30m

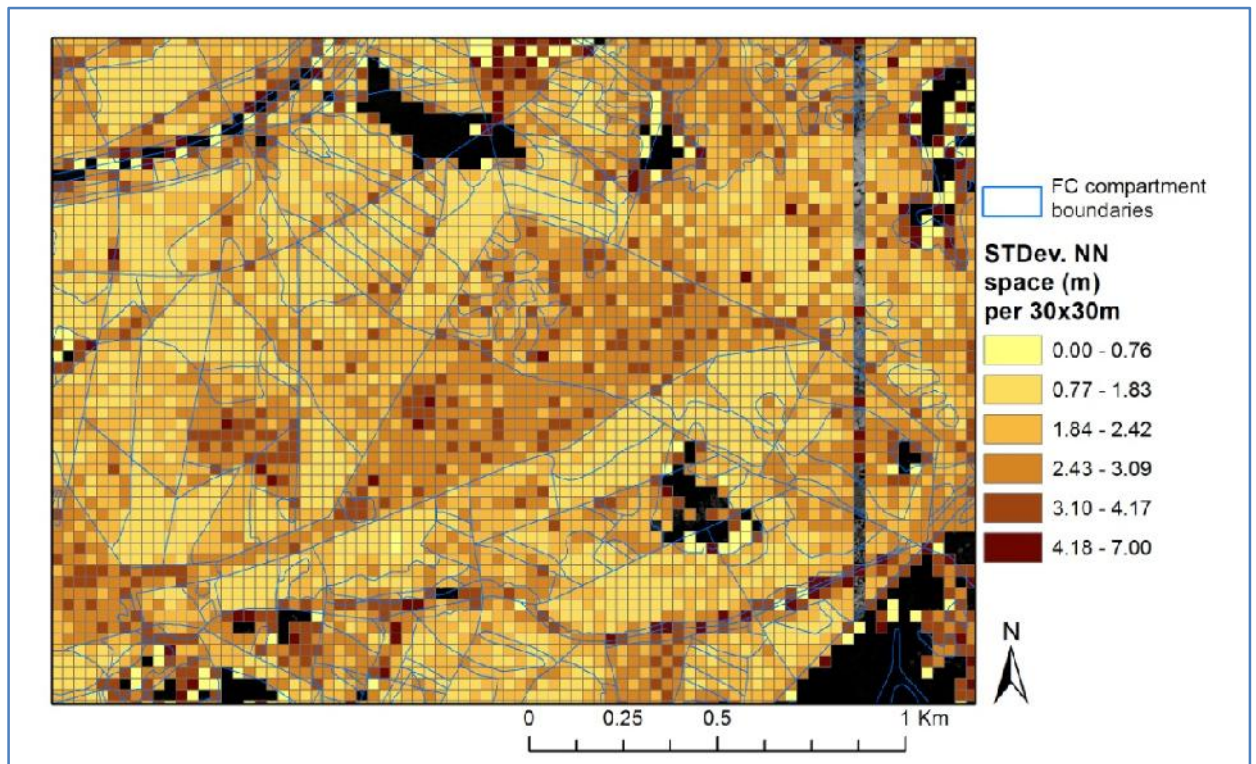


Figure 9.7 – The predicted standard deviation of tree stem nearest neighbour spacing (m) per 30x30m

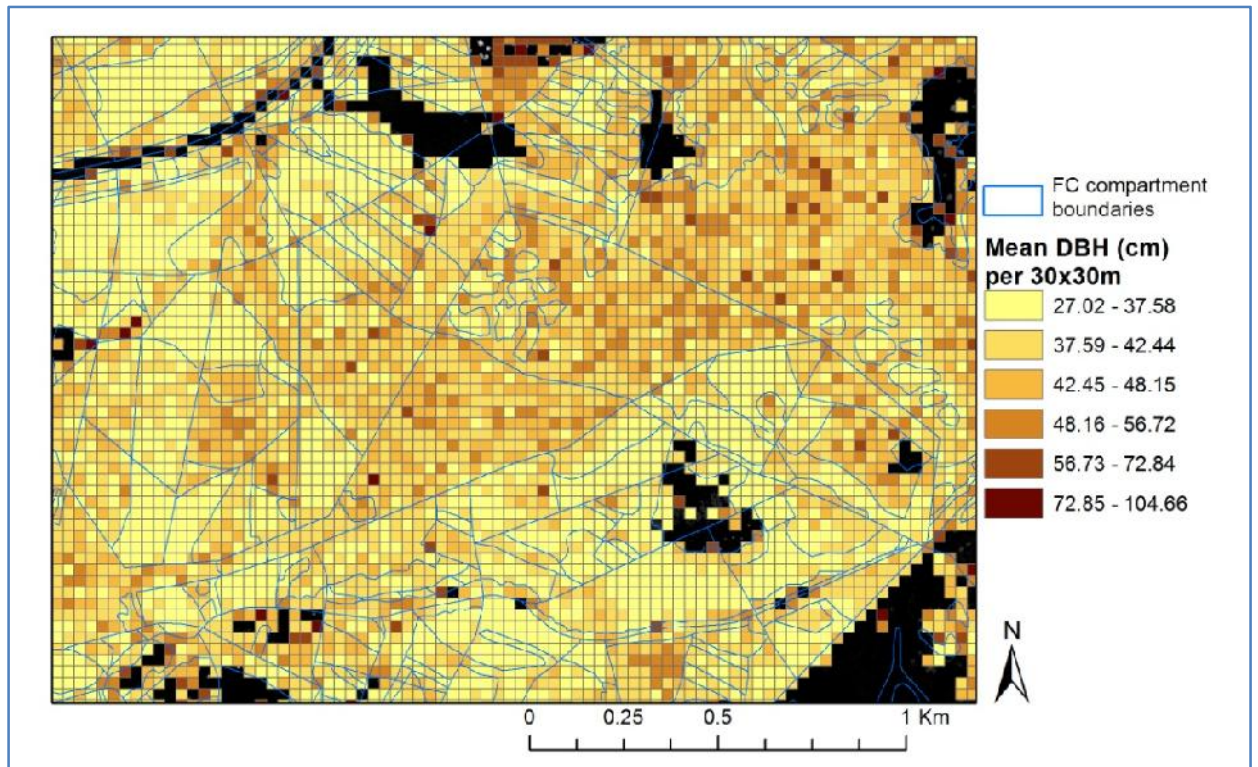


Figure 9.8 – The predicted average DBH (cm) of tree stems per 30x30m

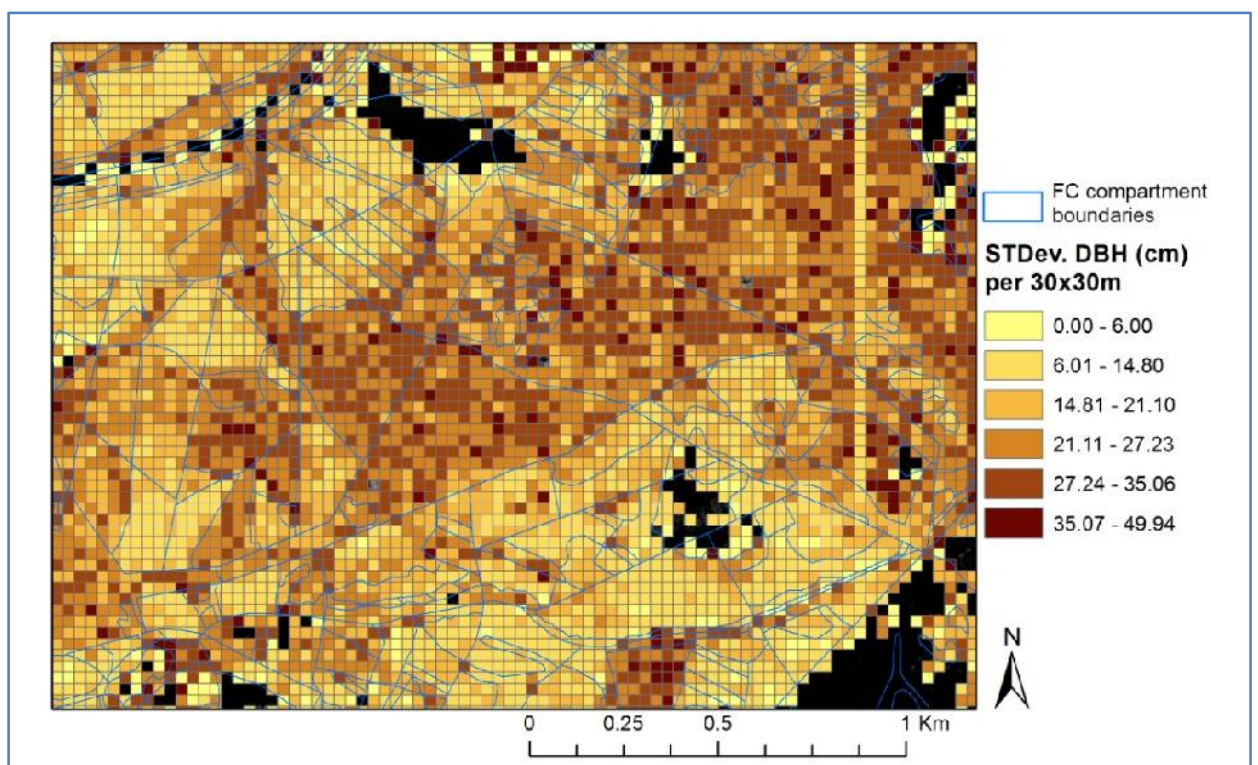


Figure 9.9 – The predicted standard deviation of DBH (cm) of tree stems per 30x30m

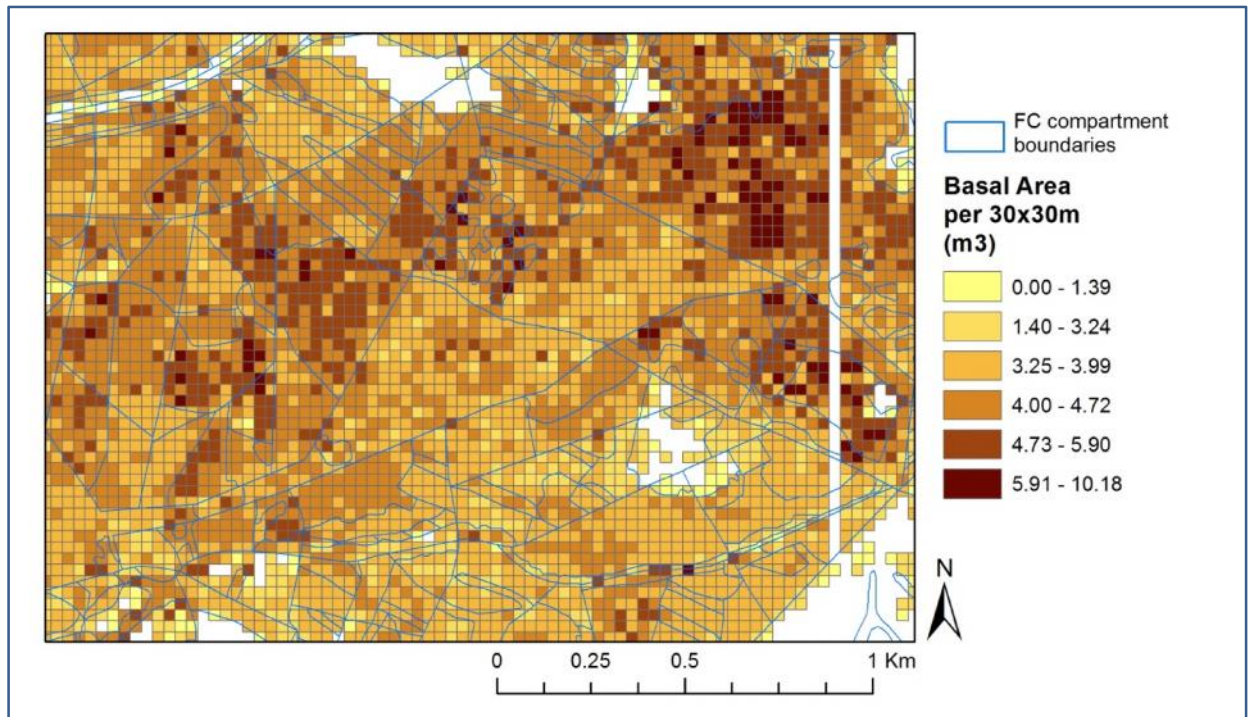


Figure 9.10 – The predicted total basal area (m²) of tree stems per 30x30m

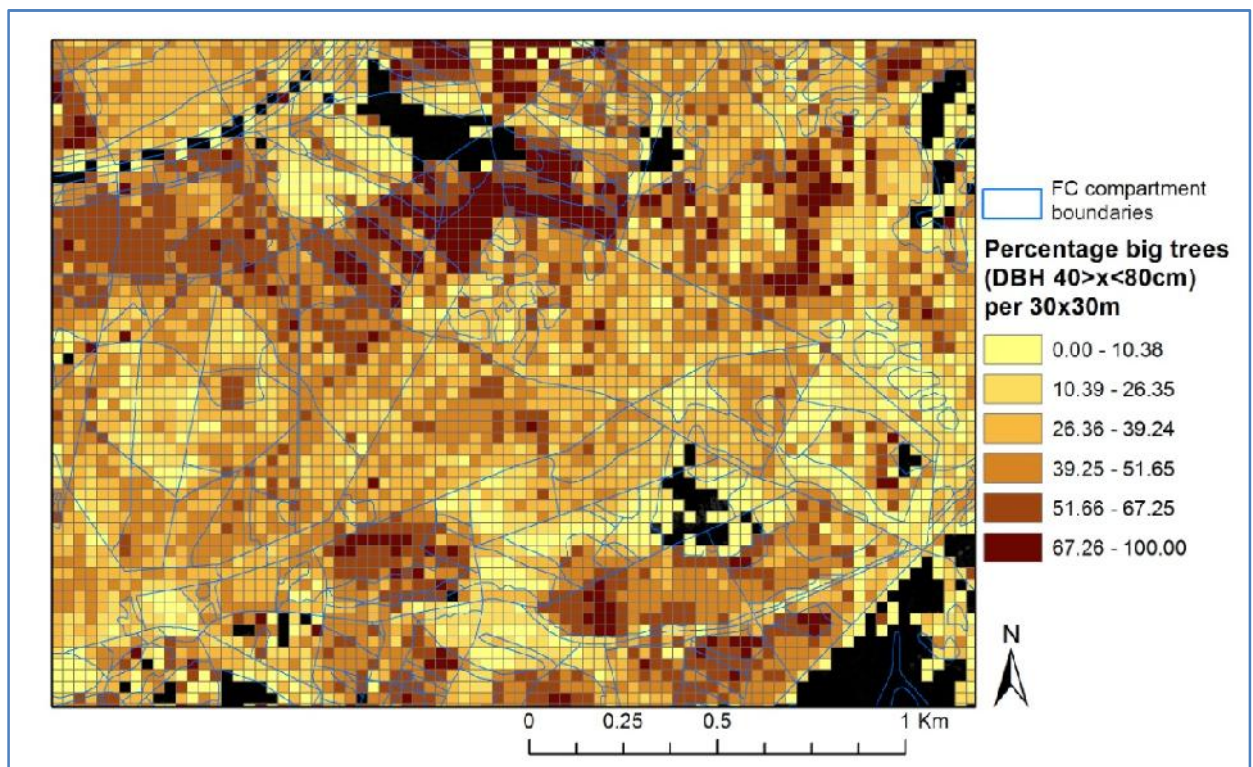


Figure 9.11 – The predicted percentage of large trees (i.e. DBH 40 > x < 80 cm) per 30x30m

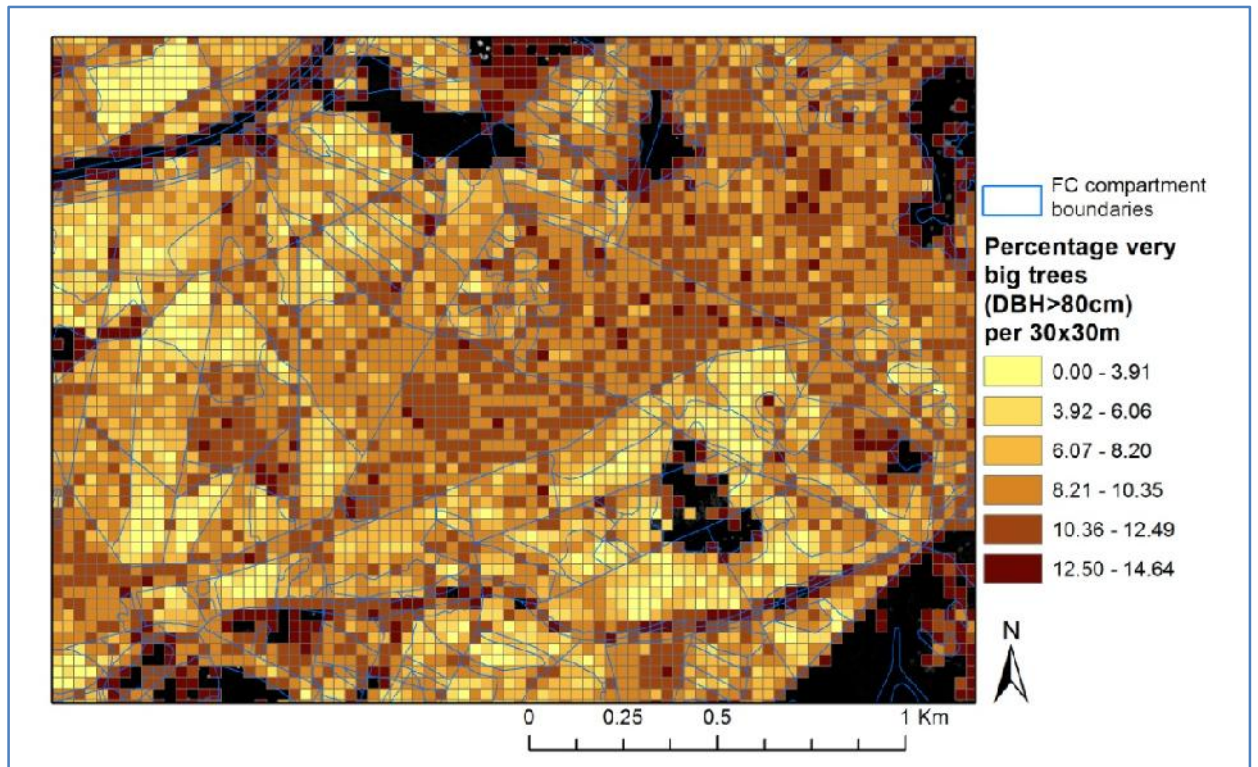


Figure 9.12 – The predicted percentage of very large trees (i.e. DBH > 80cm) per 30x30m

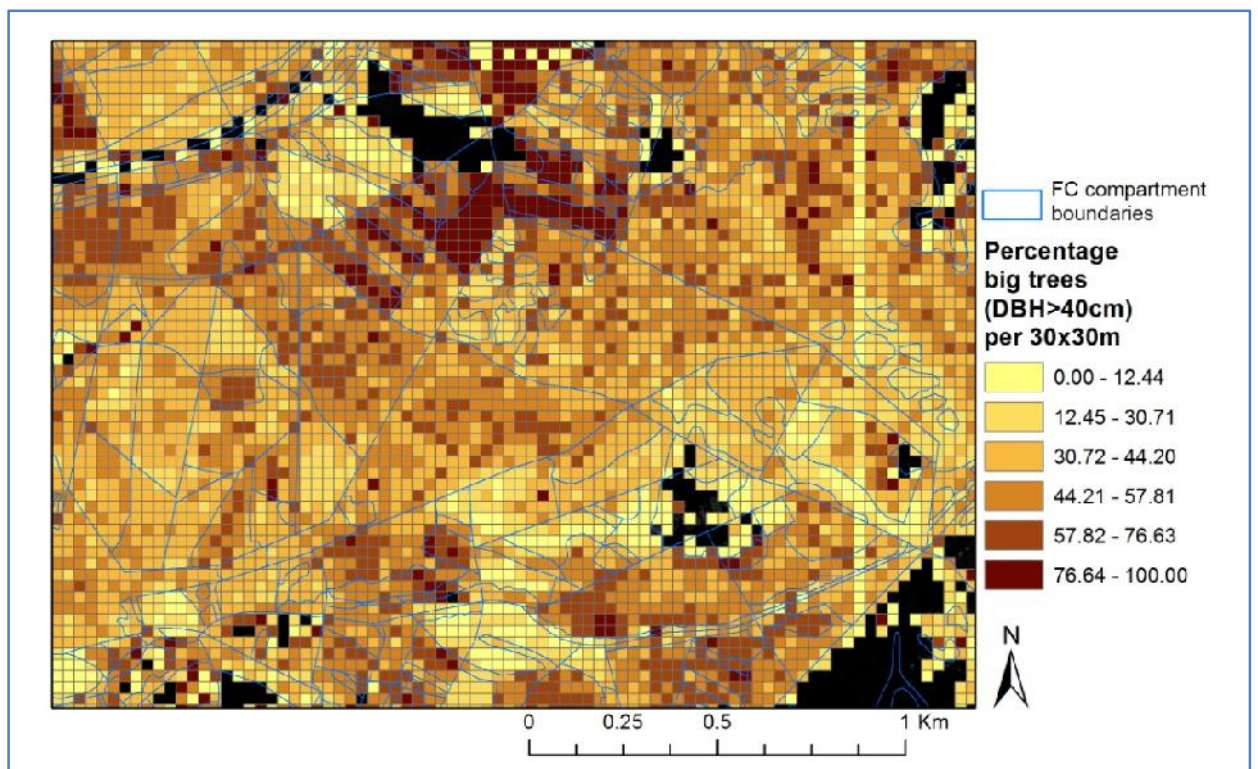


Figure 9.13 – The predicted percentage of large trees (i.e. DBH > 40cm) per 30x30m

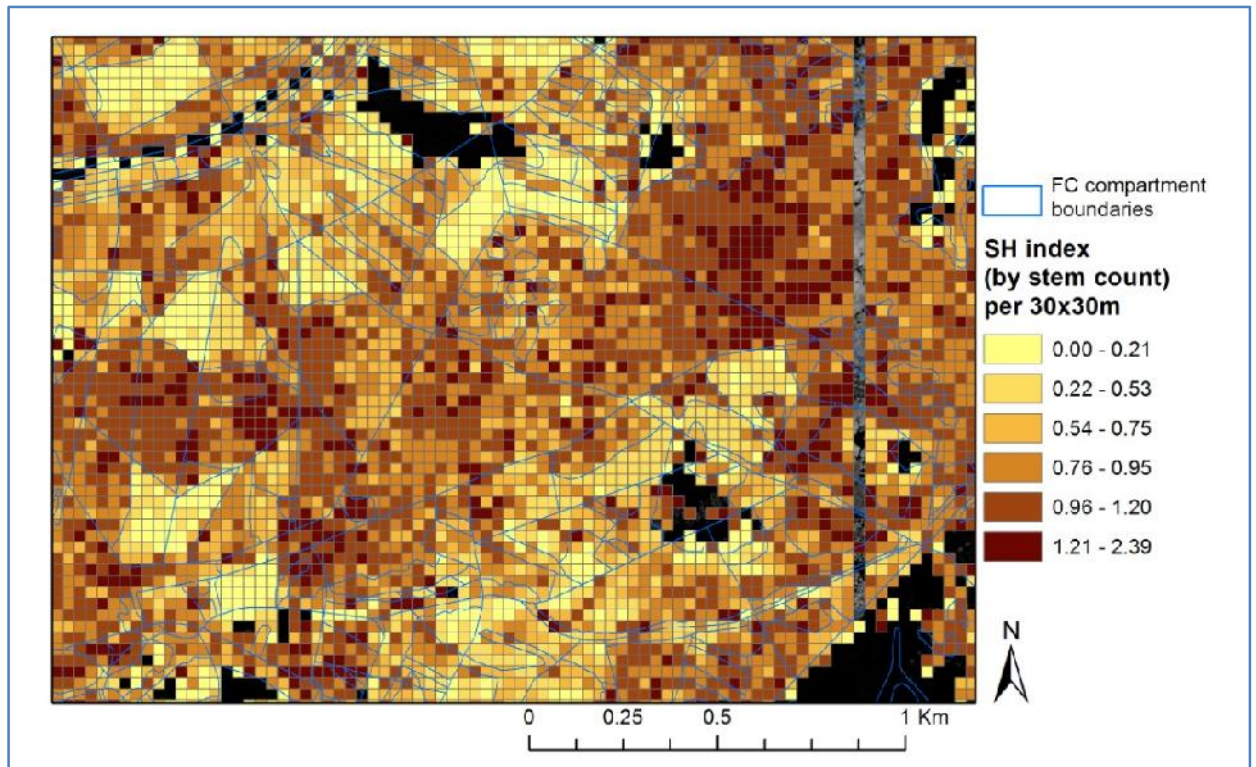


Figure 9.14 – The predicted SH index value (for species counts per stem) per 30x30m

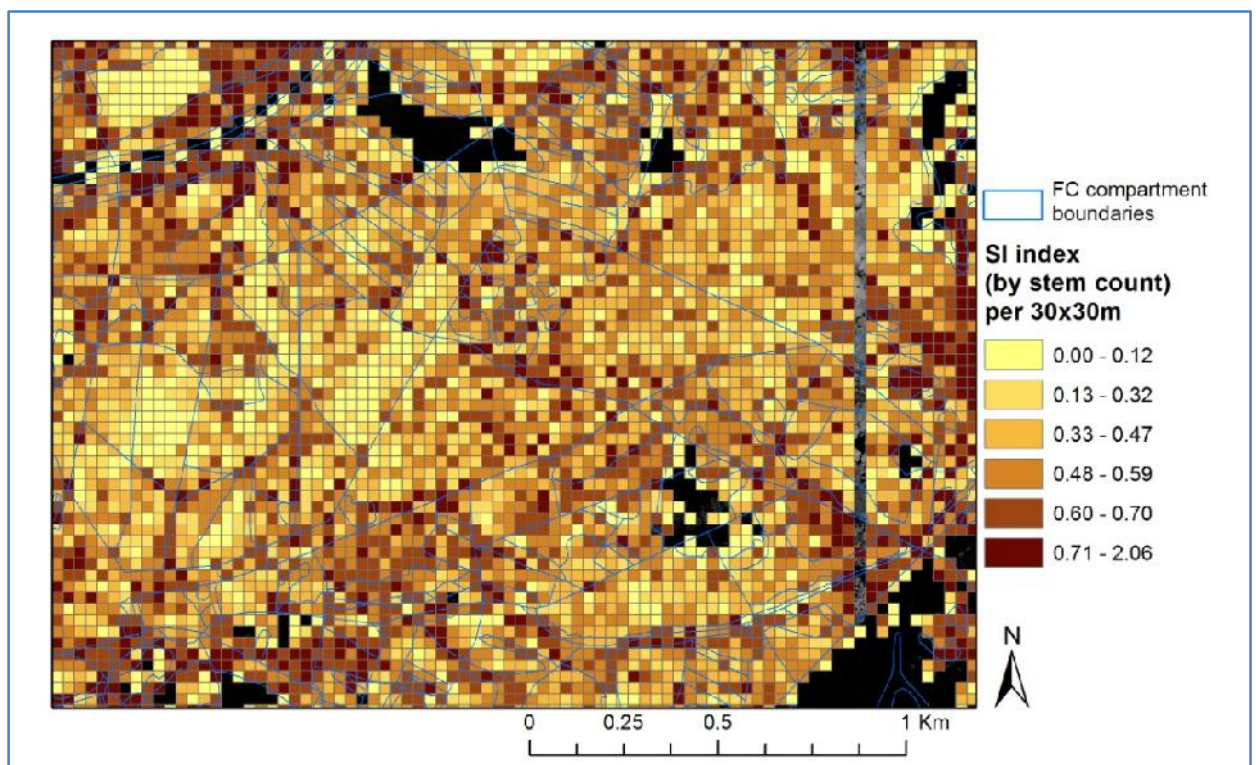


Figure 9.15 – The predicted SI index value (for species counts per stem) per 30x30m

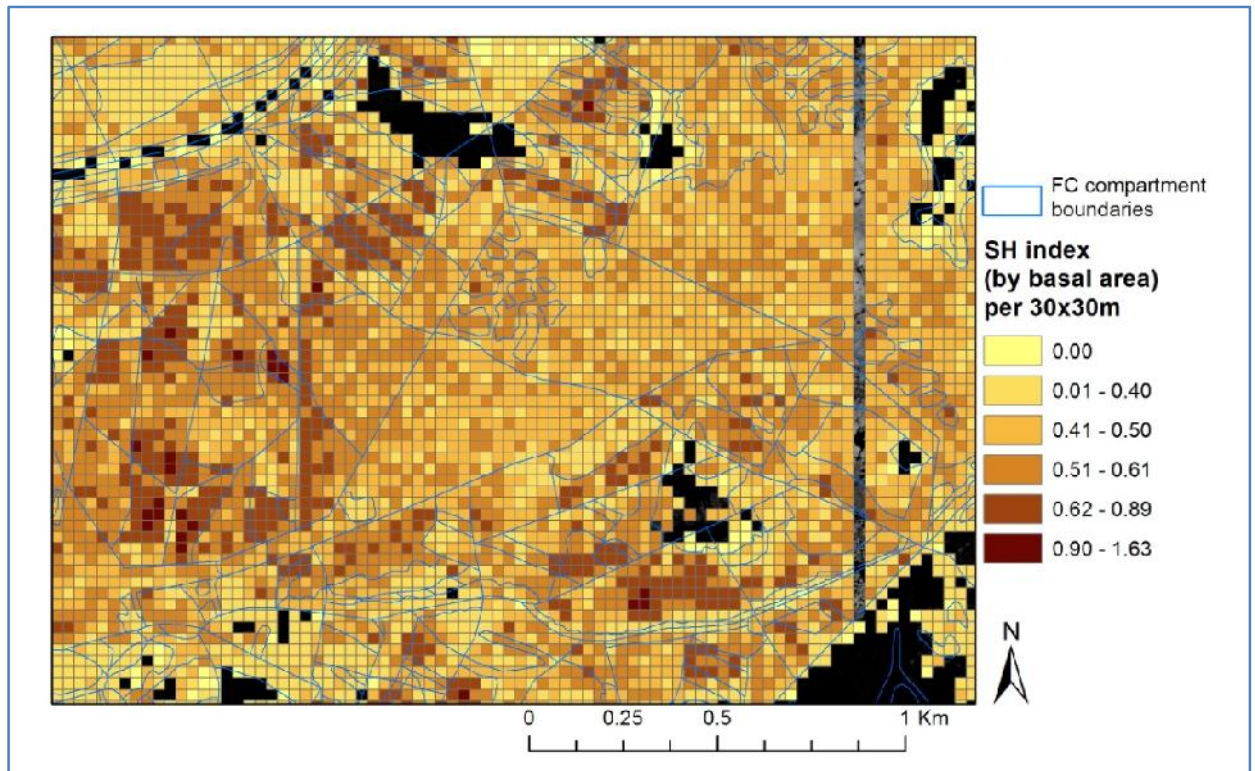


Figure 9.16 – The predicted SH index value (for species proportion defined by basal area) per 30x30m

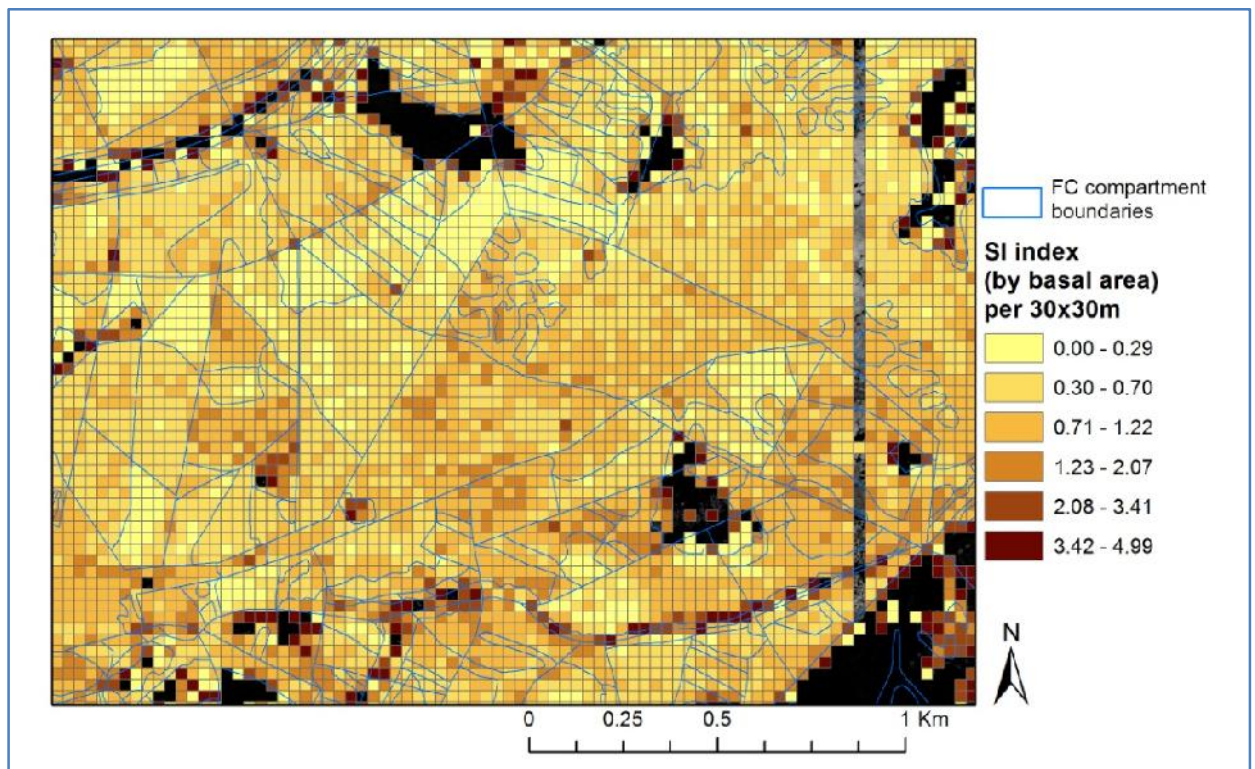


Figure 9.17 – The predicted SI index value (for species proportion defined by basal area) per 30x30m

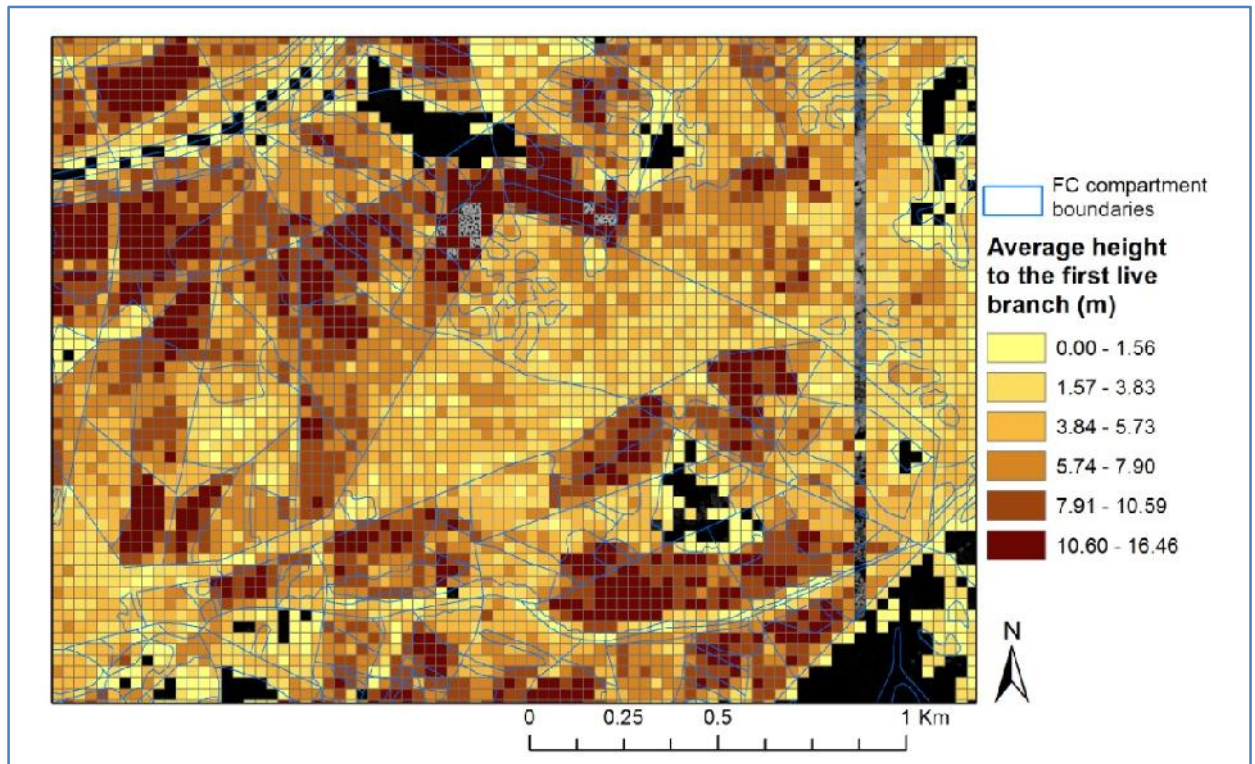


Figure 9.18 – The predicted average height of the first live branch (m) per 30x30m

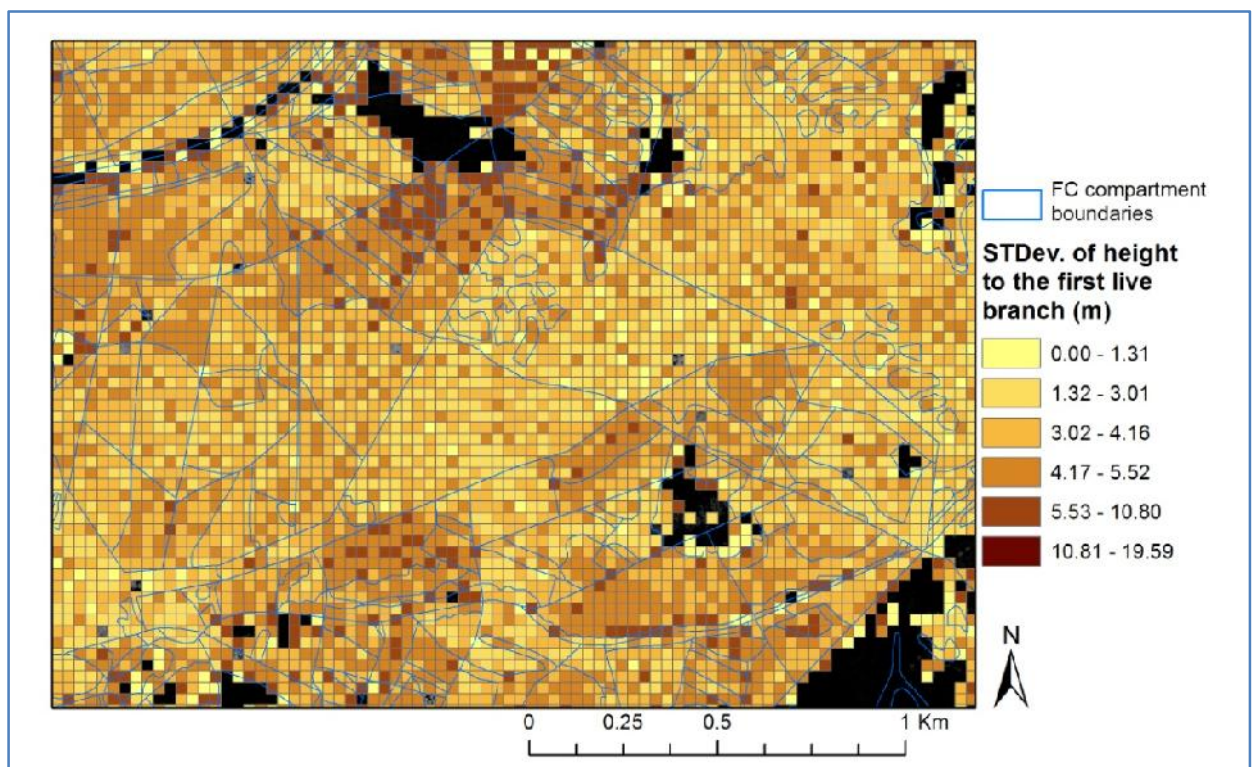


Figure 9.19 – The predicted standard deviation of height of the first live branch (m) per 30x30m

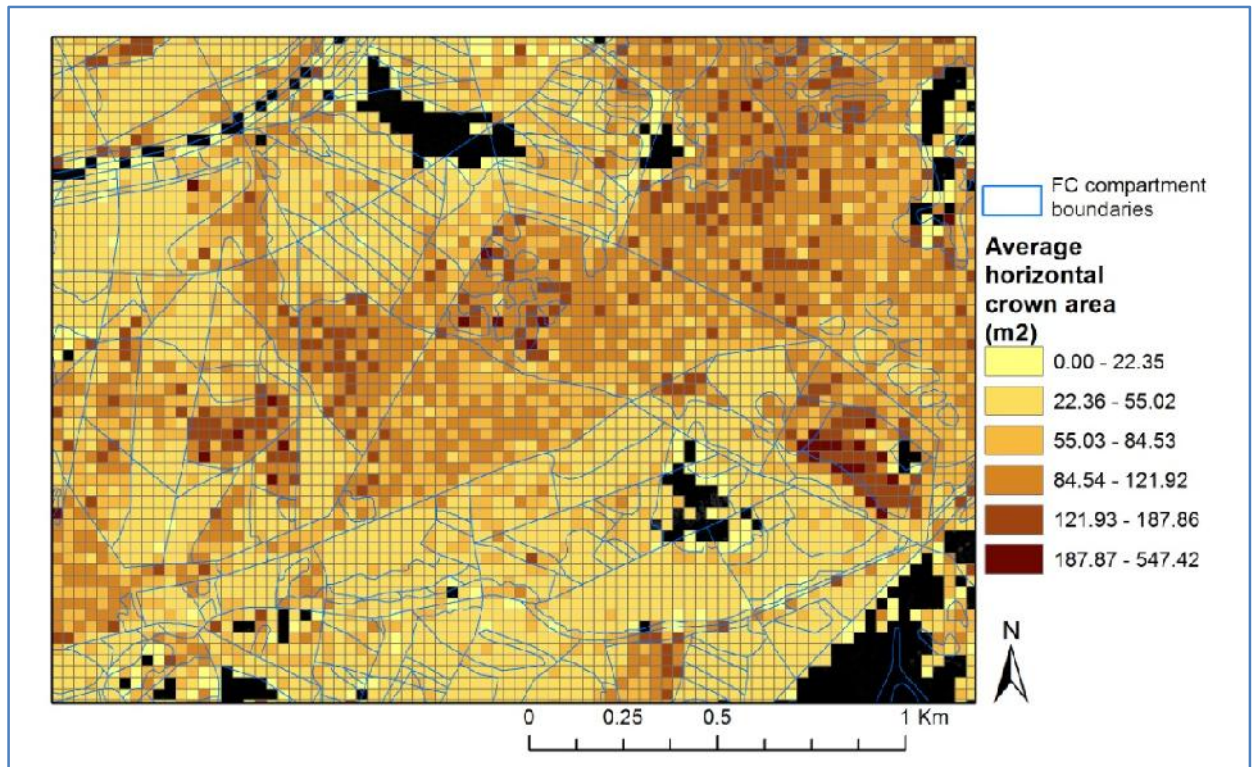


Figure 9.20 – The predicted average horizontal crown area (m²) per 30x30m

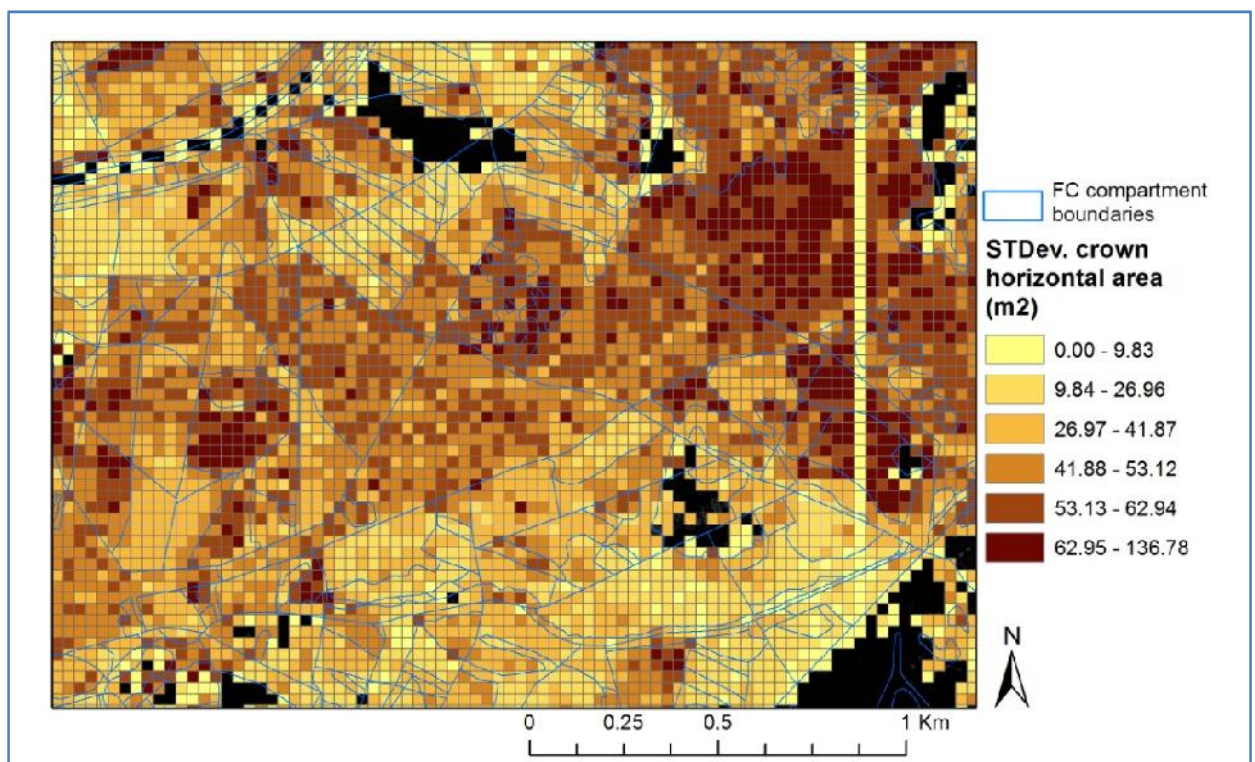


Figure 9.21 – The predicted standard deviation of horizontal crown area (m²) per 30x30m

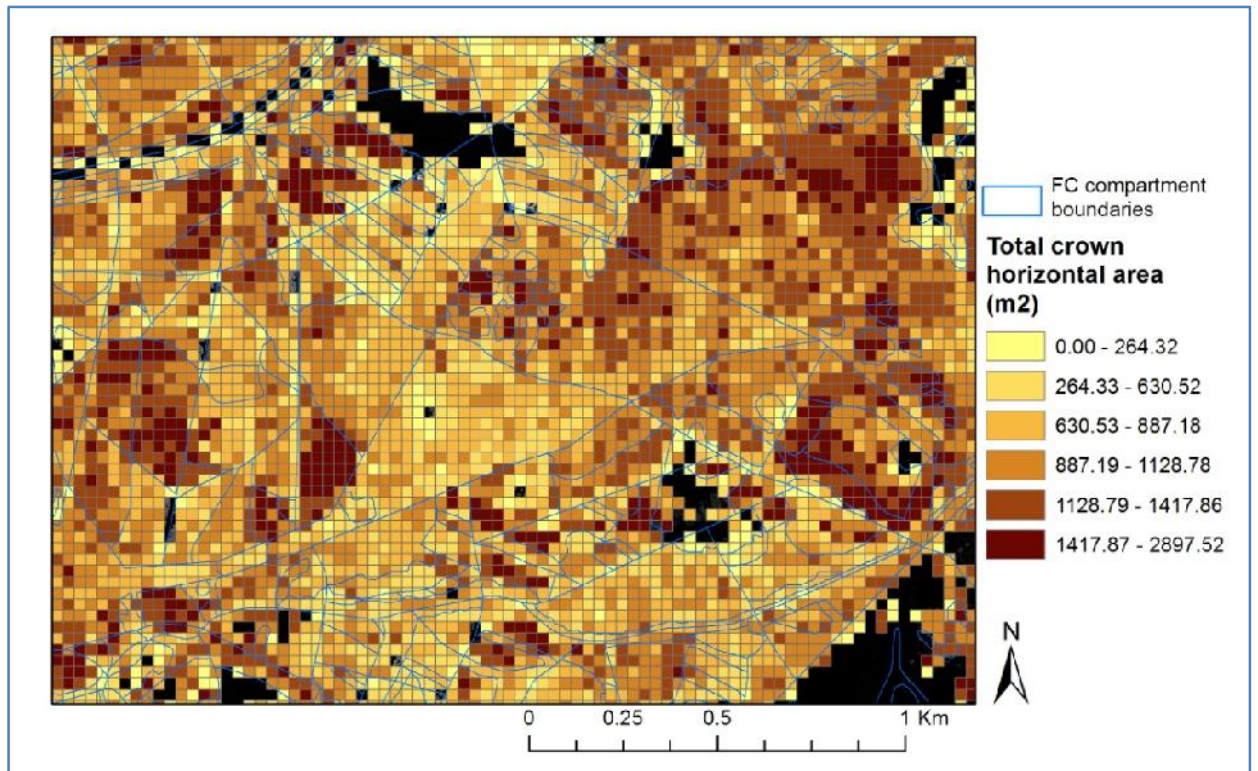


Figure 9.22 – The predicted total horizontal crown area (m²) per 30x30m

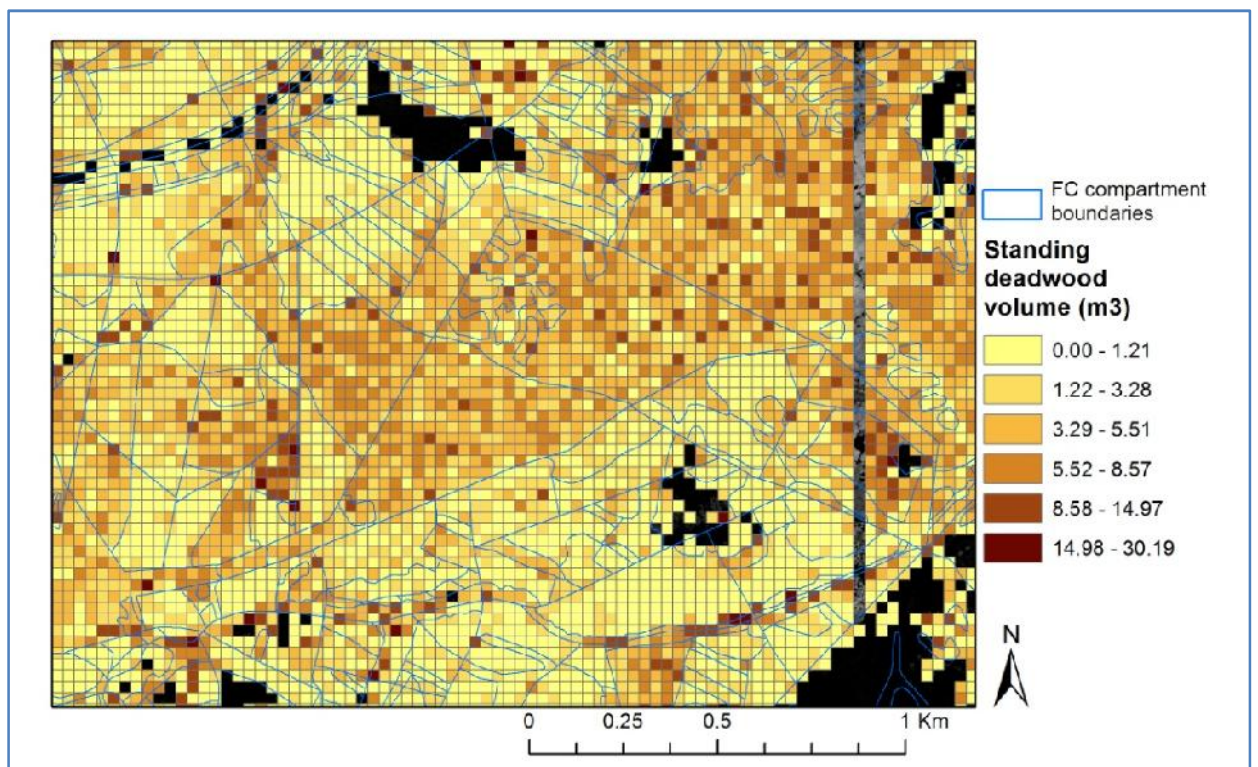


Figure 9.23 – The predicted total of standing deadwood volume (m³) per 30x30m

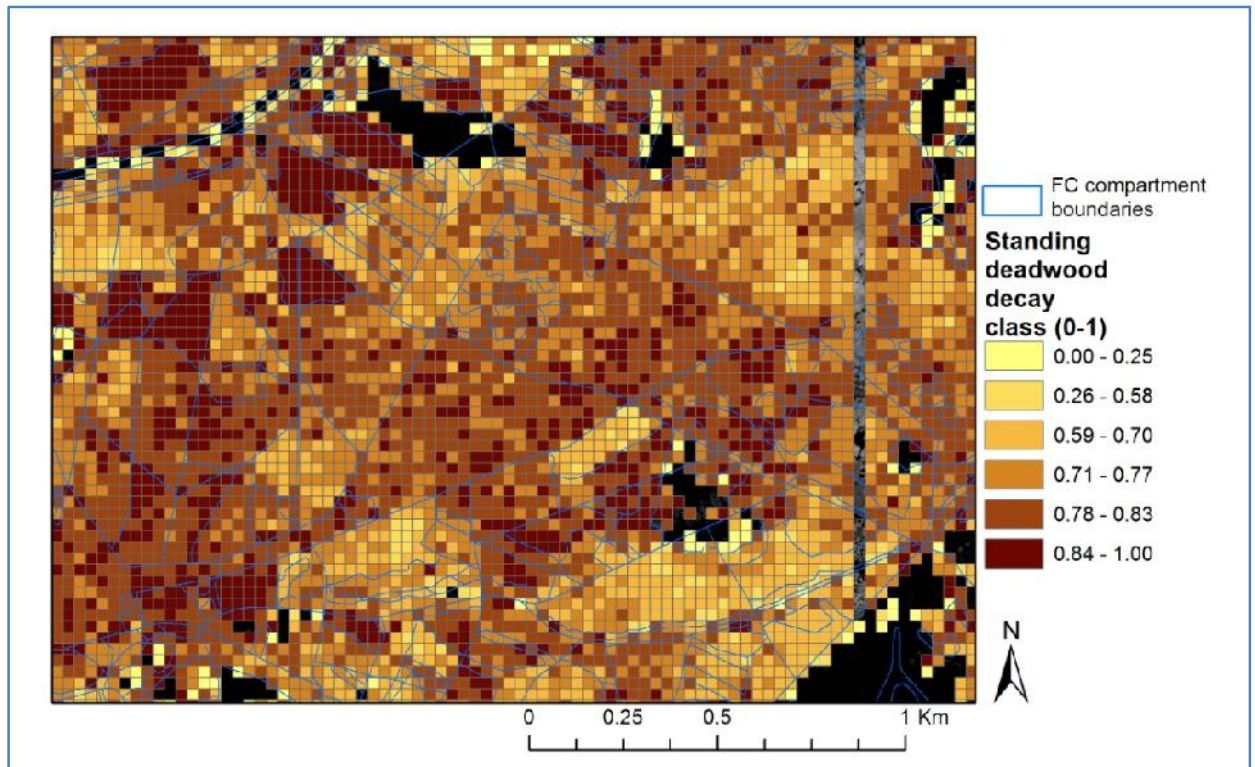


Figure 9.24 – The predicted decay class of standing deadwood per 30x30m

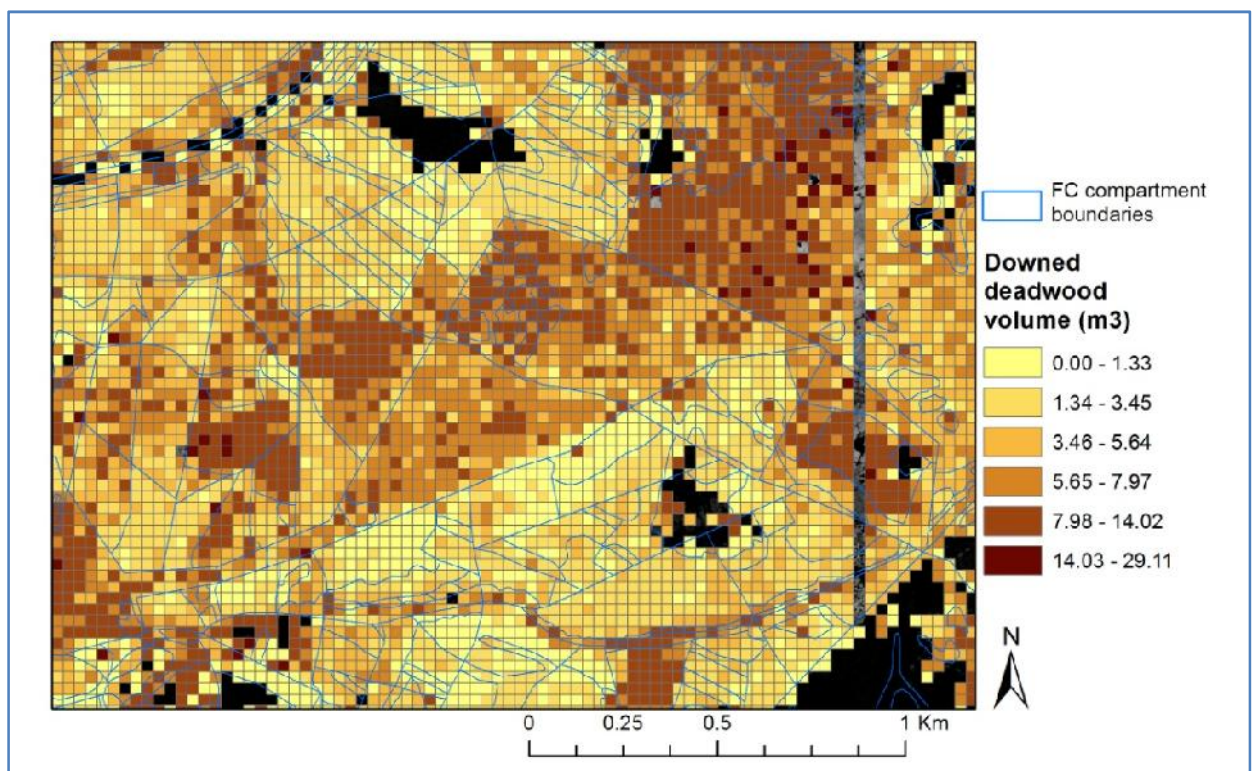


Figure 9.25 – The predicted total of downed deadwood volume (m³) per 30x30m

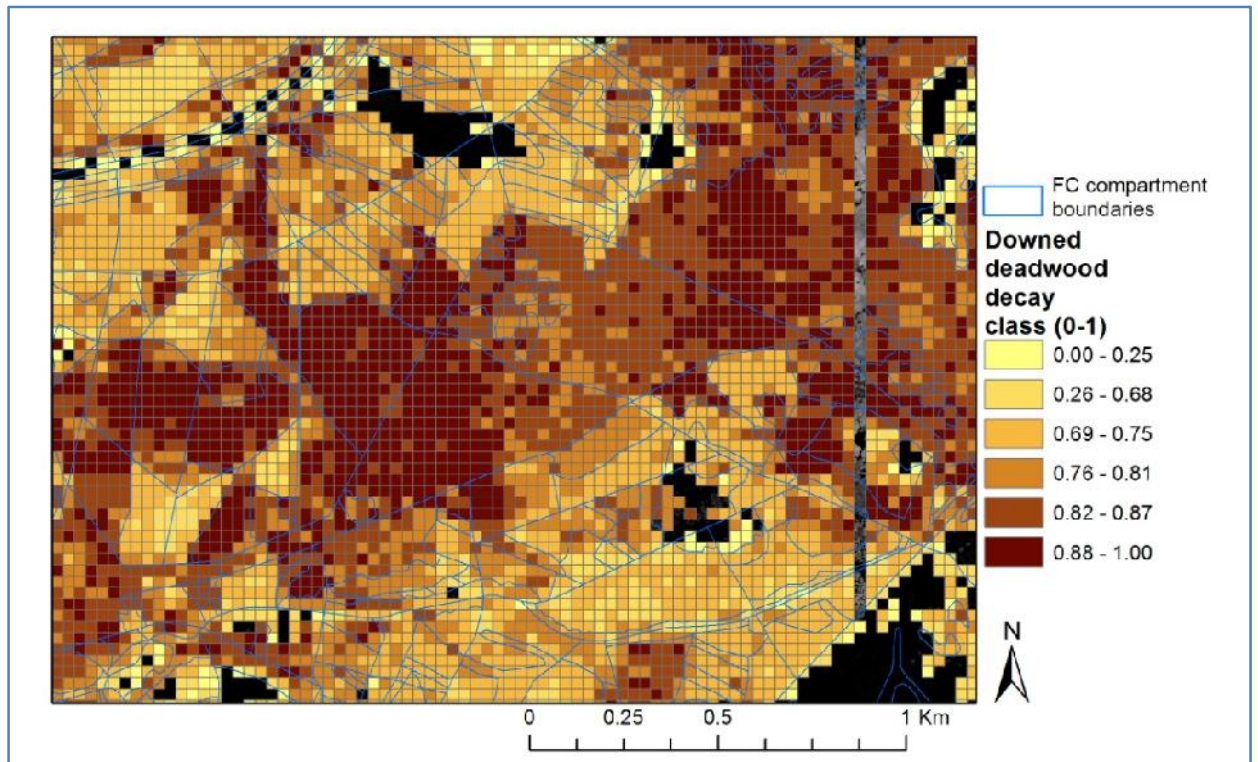


Figure 9.26 – The predicted decay class of standing deadwood per 30x30m

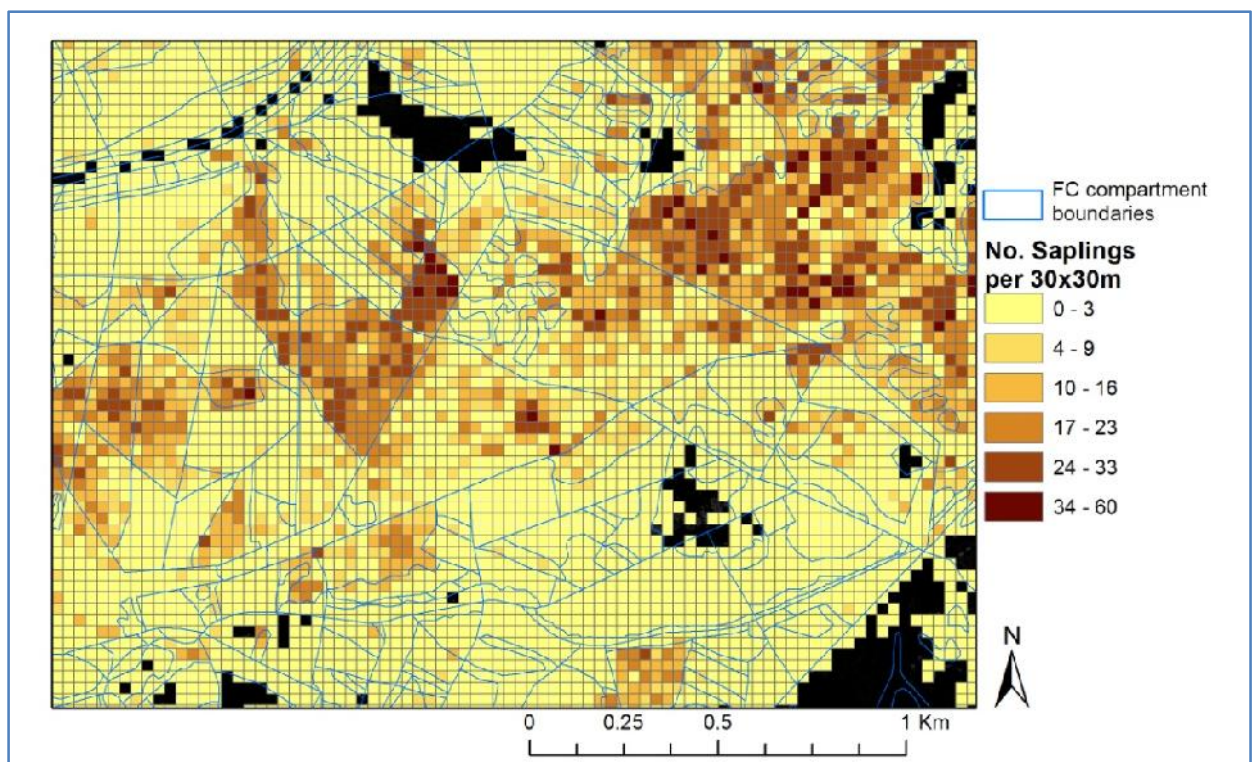


Figure 9.27 – The predicted number of saplings per 30x30m

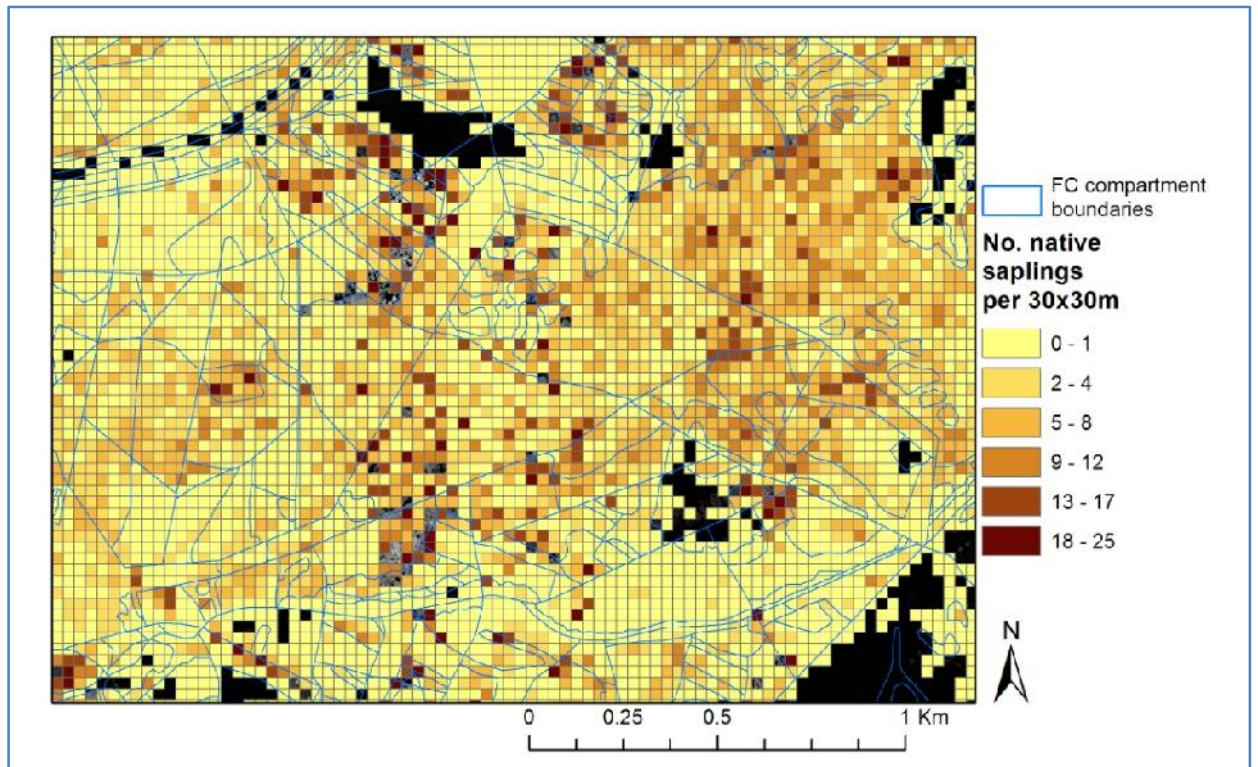


Figure 9.28 – The predicted number of saplings of native species per 30x30m

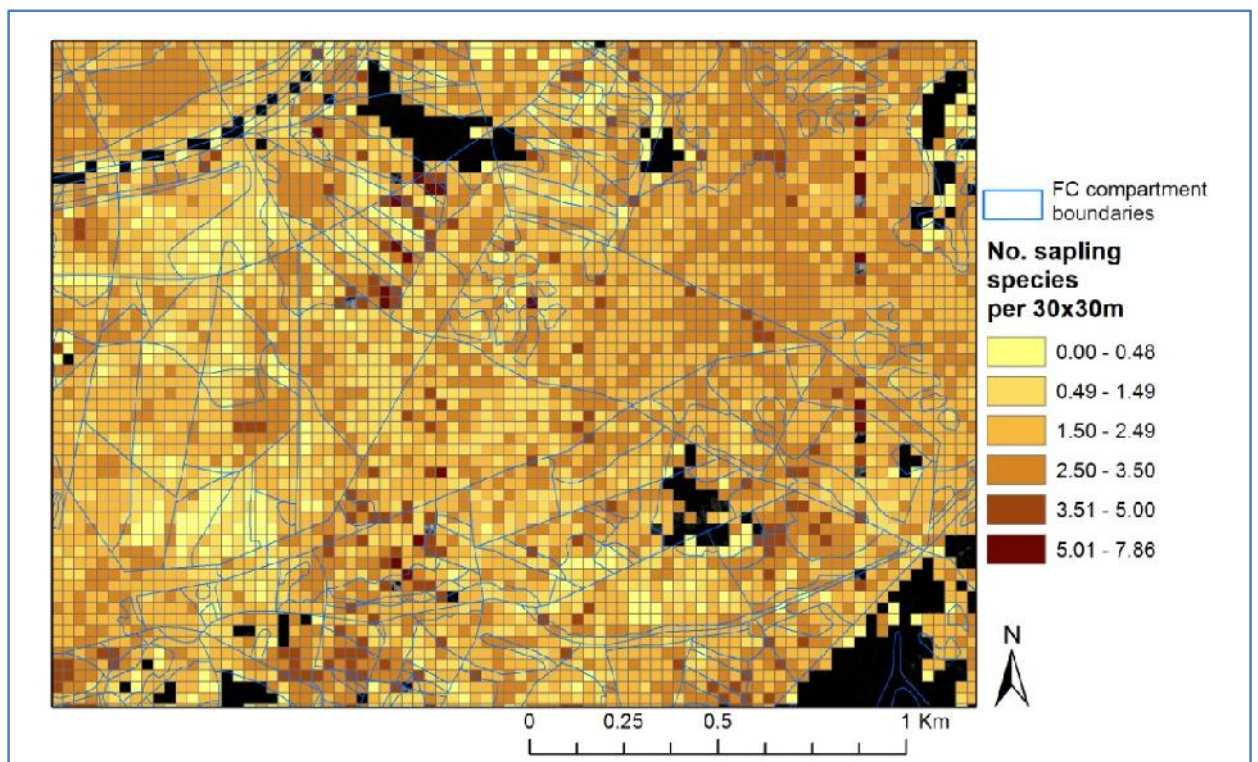


Figure 9.29 – The predicted number of saplings species per 30x30m

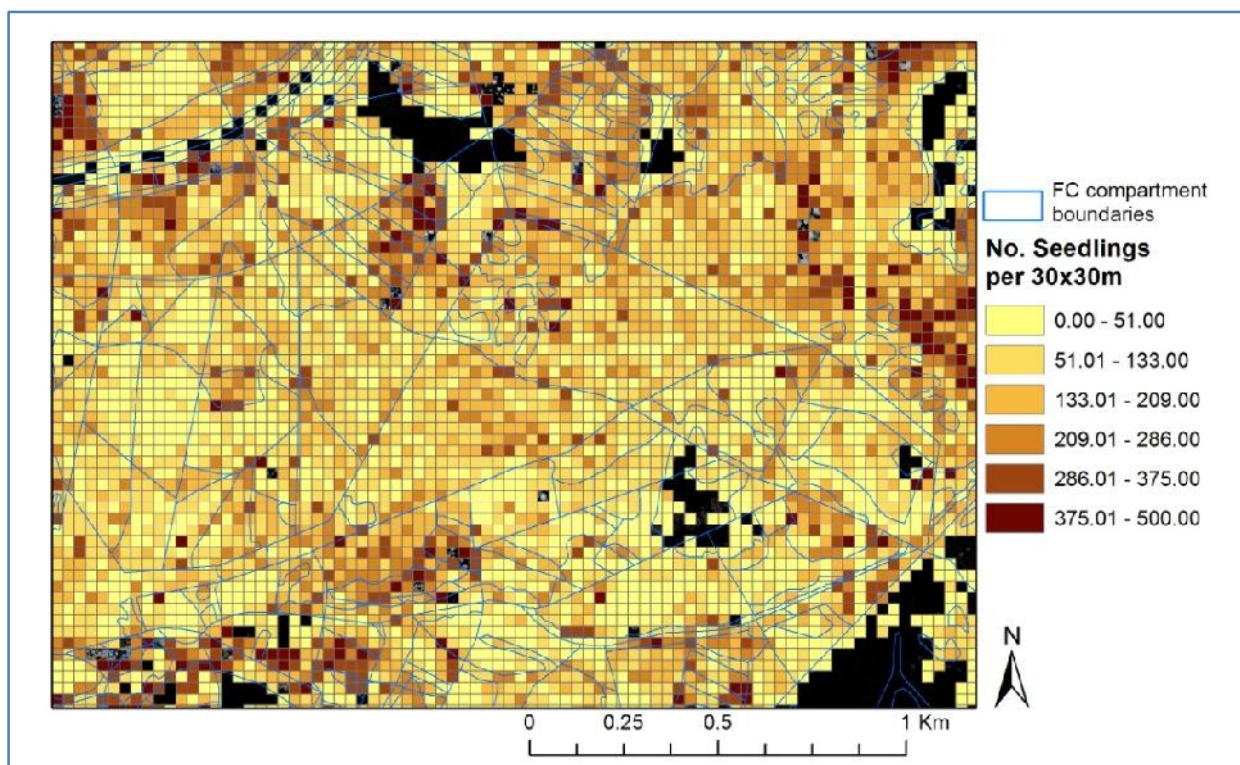


Figure 9.30 – The predicted number of seedlings per 30x30m

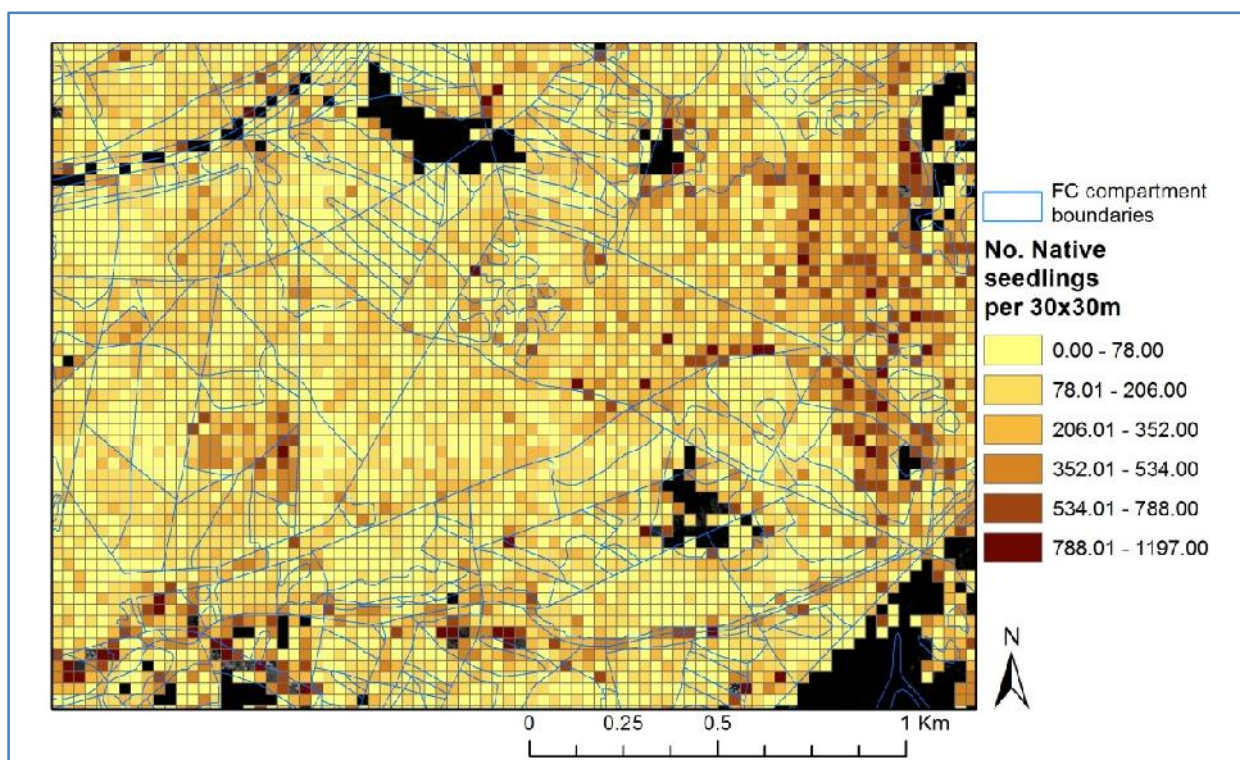


Figure 9.31 – The predicted number of seedlings of native species per 30x30m

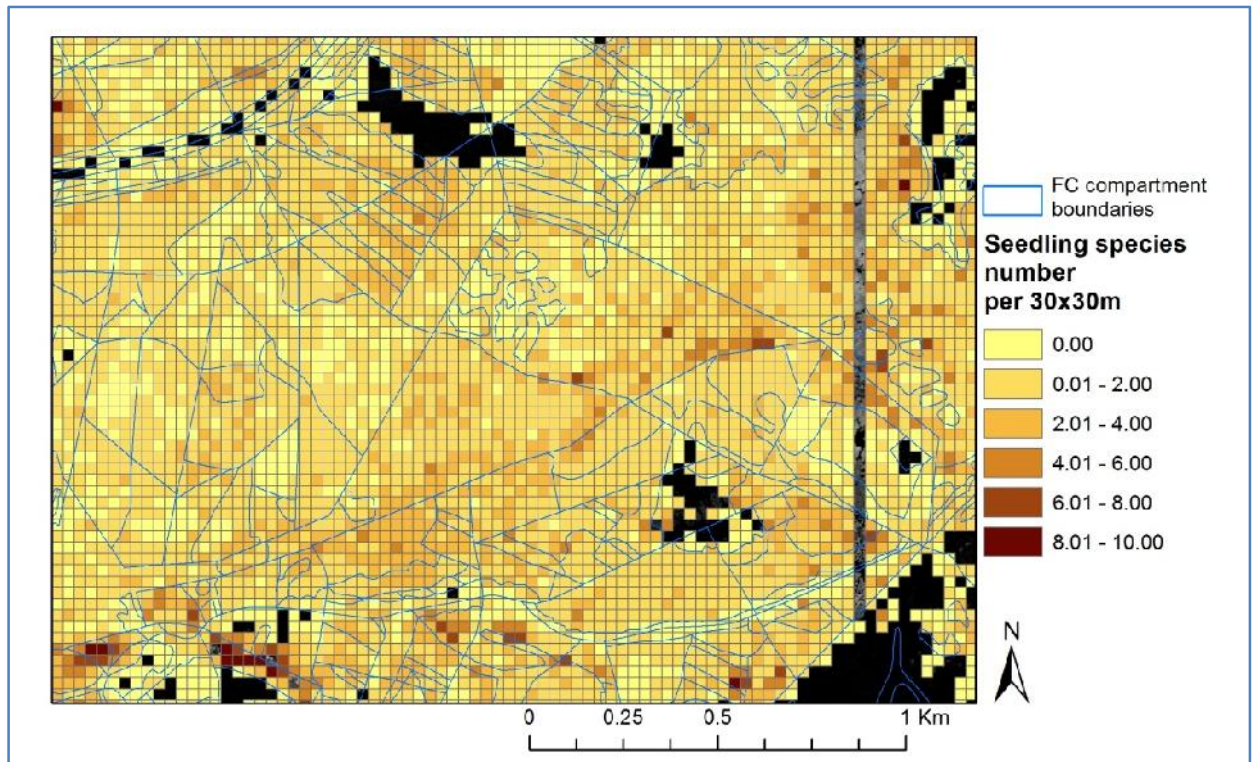


Figure 9.32 – The predicted of seedlings species per 30x30m

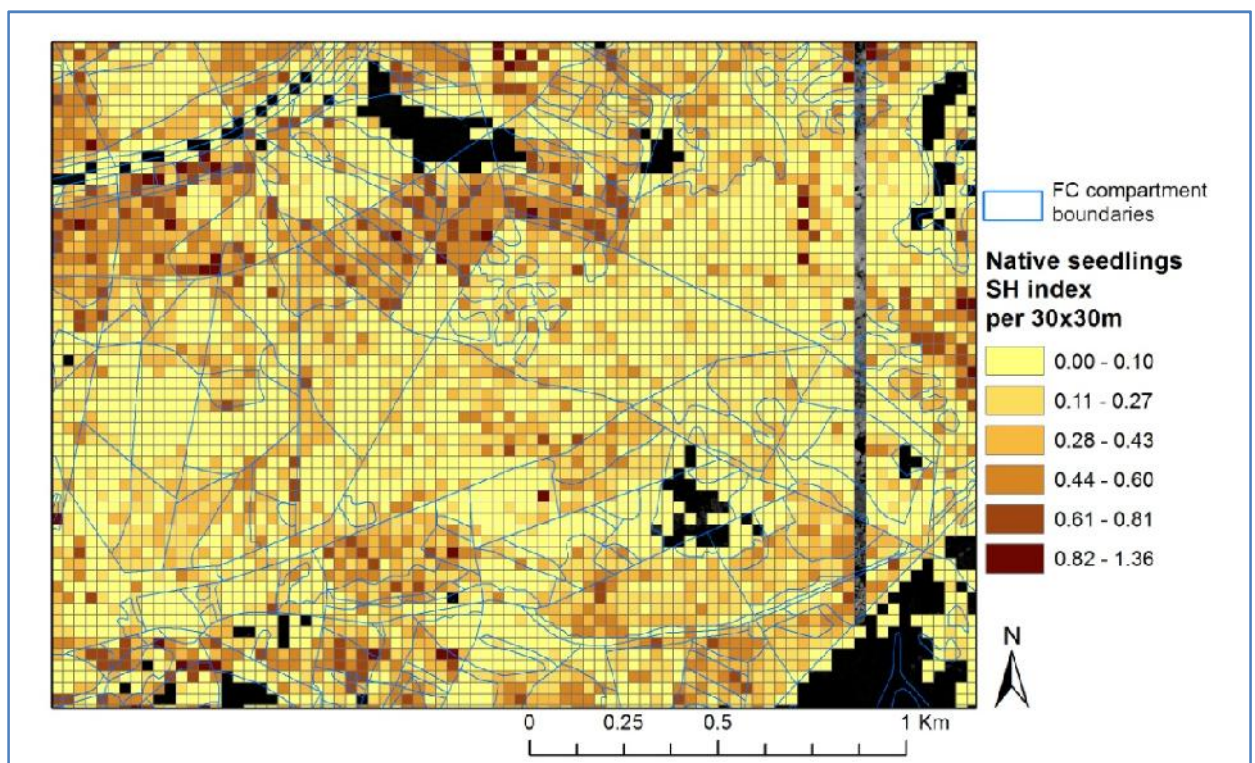


Figure 9.33 – The predicted of SH index for native seedlings species per 30x30m

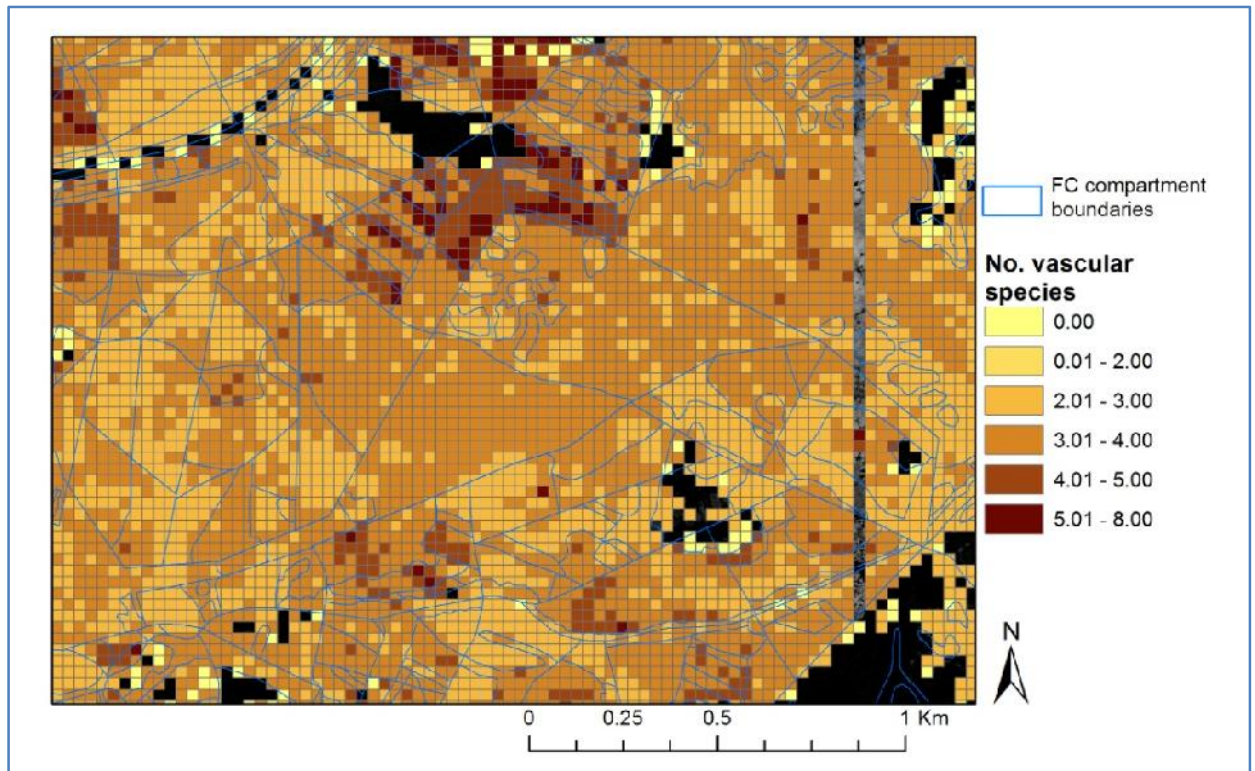


Figure 9.34 – The predicted number of vascular species per 30x30m

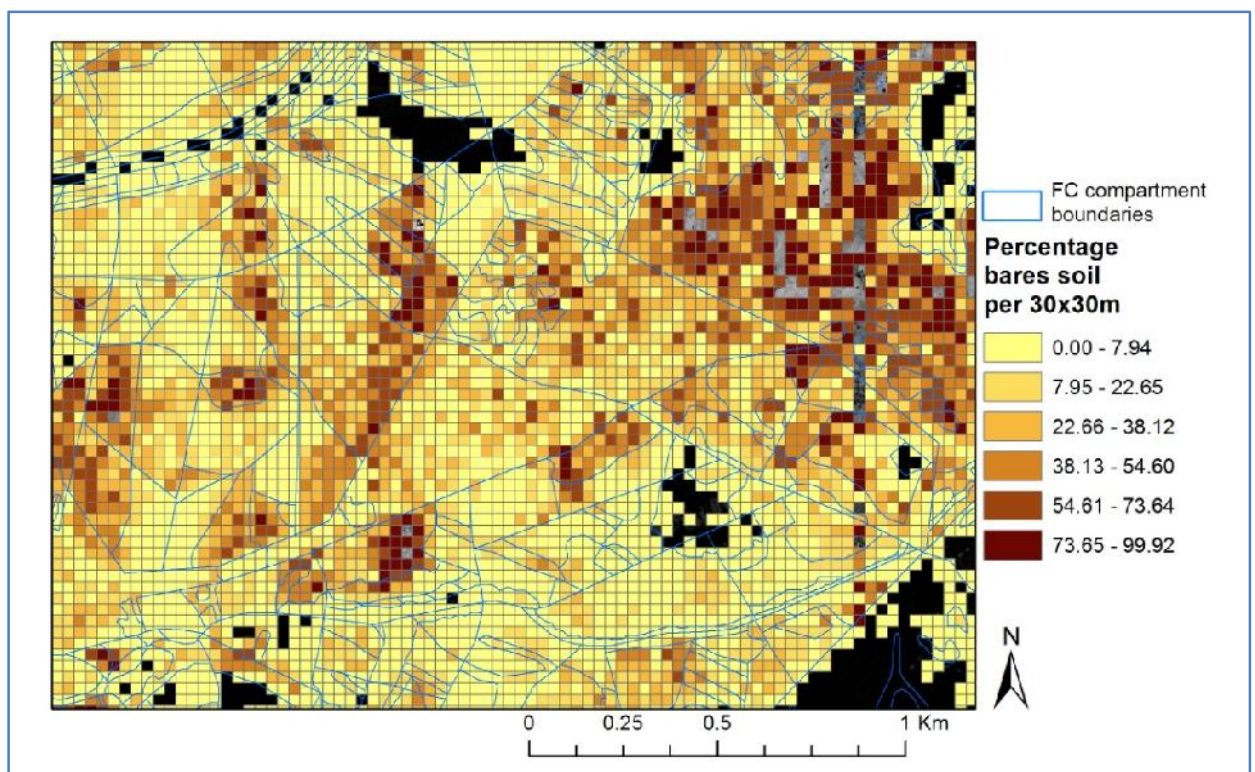


Figure 9.35 – The predicted percentage of bare earth/soil per 30x30m

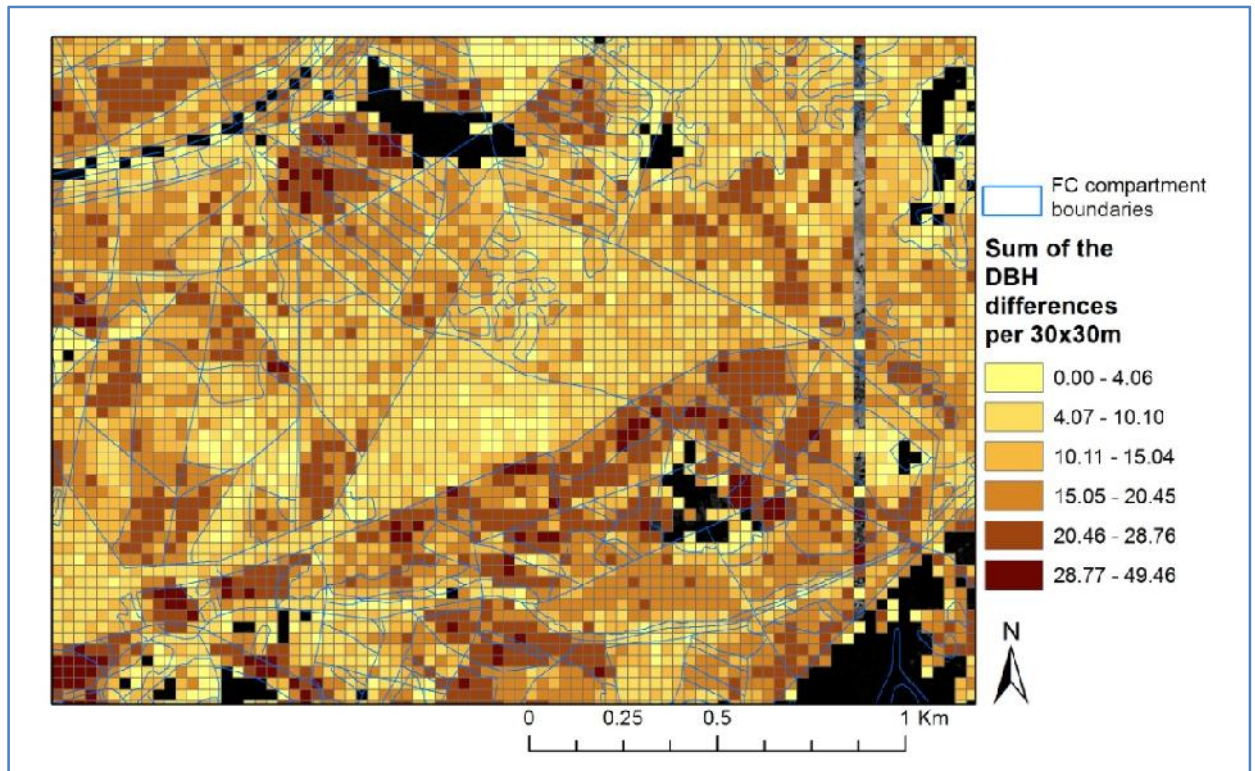


Figure 9.36 – The predicted difference of the sum of DBH (cm) differences per 30x30m

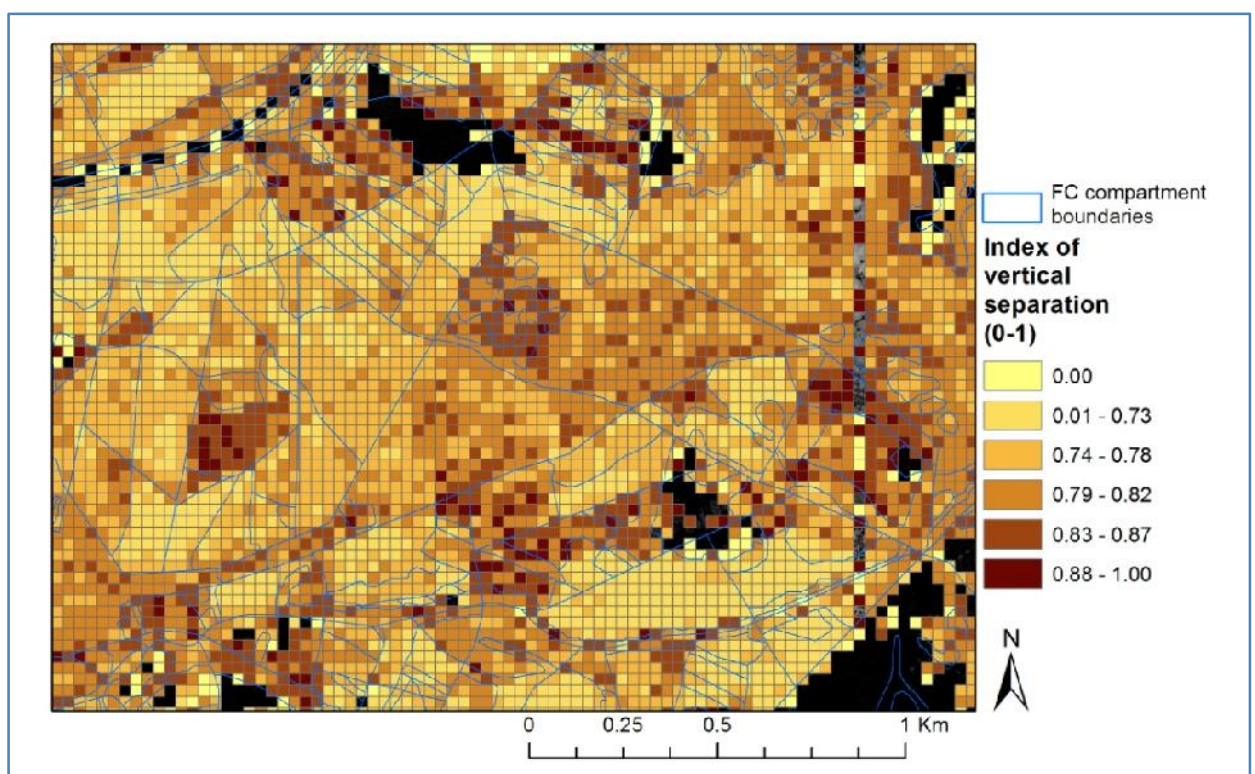


Figure 9.37 – The predicted Index of vertical separation per 30x30m

9.2 Condition index mapping

Six condition indices were calculated from the various remote sensing derived products relating to summarised ITC and statistically derived field plot-level metrics. The first two condition index calculations reported in this section utilise ITC metrics only however. The statistically derived metrics reported here conform to the best models, or more specifically those models with the smallest RMSE/NRMSE from the following dataset metrics: (i) hyperspectral; (ii) DR LiDAR, (iii) FW LiDAR; (iv) combined hyperspectral and DR LiDAR; and (v) combined hyperspectral and FW LiDAR, as identified in Chapter 8.4.

Figure 9.38 illustrates the sample area for all the illustrations of condition indices that are mapped across the study site. This image contains the FC compartment boundaries with the primary species class, as identified by the FC, overlaid upon a 1x1m CHM as a general indication of forest structure and composition at that position. The area centres upon the semi-ancient Frame Wood, with Frame-Heath and Denny-Lodge inclosures to the north, Tantany Wood to the north-east and Hawkhill inclosure to the south.

A comparison of the condition indices calculated from the derived field-level metrics from (i) DR LiDAR, hyperspectral or DR and hyperspectral composite; or (ii) FW LiDAR, hyperspectral or FW and hyperspectral composite models is available in Appendix F section F.2. It should be noted that Figures 9.39-9.44 may contain two image artefacts due to calculation errors which could not be corrected in the available time. These mapping artefacts appear as two lines of grid-cells from north to south in the east of the map images. These errors were caused by problems in mosaicing FW LiDAR metrics where shapefile grid cells were duplicated.

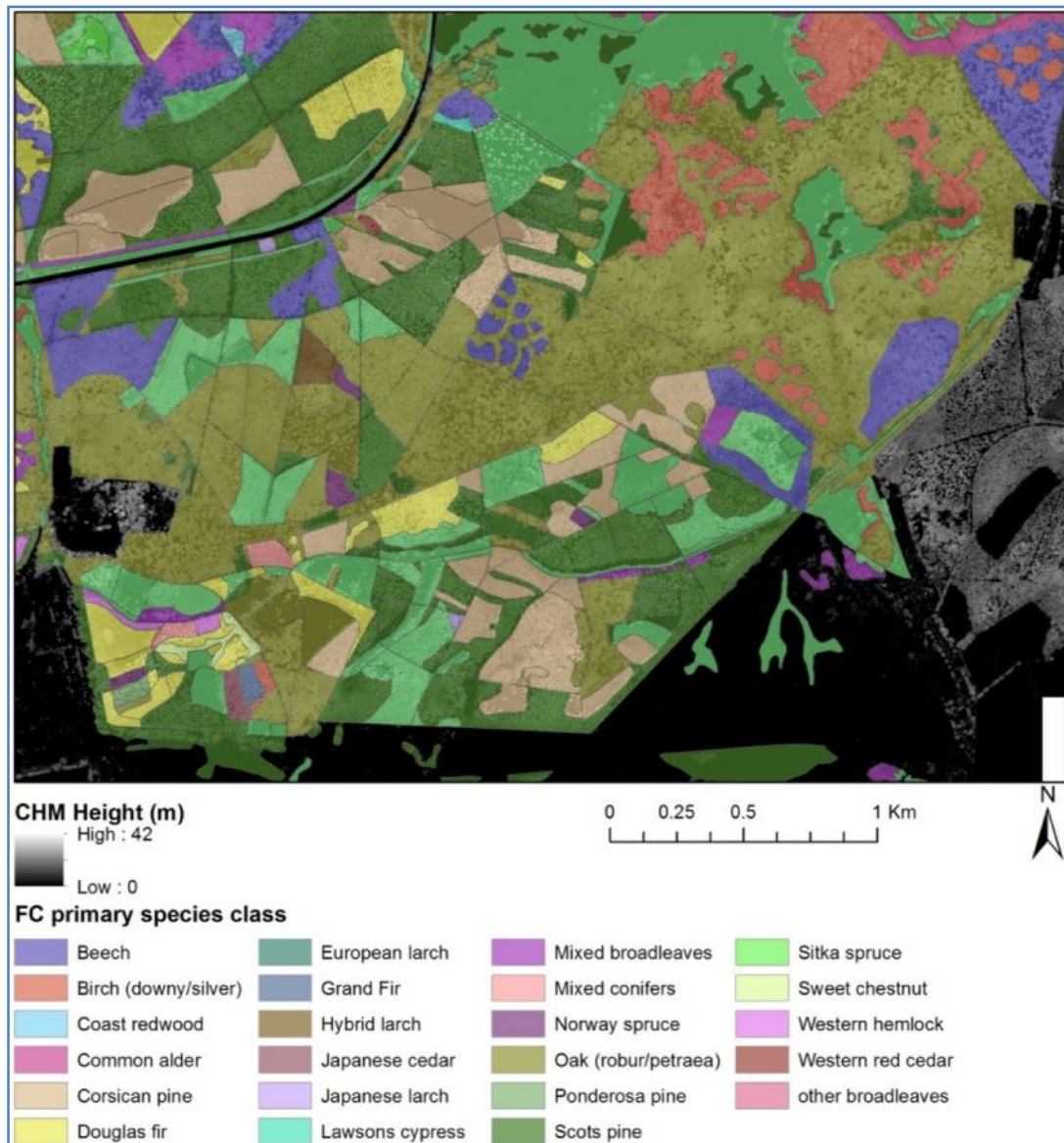


Figure 9.38 – FC compartment boundaries and primary species class overlaid upon a 1x1m CHM

9.2.1 Vertical evenness (VE) index

The VE index relates to the relative distribution of tree crown horizontal cover at different heights within the fieldwork plot. To reiterate the outputs from DR LiDAR ITC analysis using data generated from the TIFFS software was used as input for calculating the VE index. Table 9.2 summarises the VE calculated from field and DR LiDAR data. Due to the high density and/or high crown closure of tree stems in plots 24, 30 and 38 the field VE index could not be estimated. The overall RMSE for the LiDAR-derived VE was 0.292. Eight of the remaining seventeen VE estimates were below the field estimates, where the index values were less than half of the corresponding field derived value. A number of values were returned as 0. This occurred where ITC measurements for the extracted field-plot extent recorded very similar heights. Underestimation occurred for both coniferous and deciduous dominated plots; however deciduous plots were most effected.

Table 9.2 – Comparison of field vs. remote sensing derived VE index

Site no.	Class	Field calc. VE index	DR LiDAR VE index
22	Deciduous	0.318	0.314
23	Coniferous	0.380	0.372
25	Deciduous	0.479	0.111
26	Deciduous	0.436	0.312
27	Deciduous	0.685	0.179
28	Coniferous	0.369	0.325
29	Deciduous	0.539	0.000
31	Coniferous	0.288	0.228
32	Mixed	0.396	0.289
33	Coniferous	0.386	0.054
34	Mixed	0.562	0.375
35	Coniferous	0.478	0.236
36	Coniferous	0.329	0.155
37	Deciduous	0.505	0.095
39	Coniferous	0.517	0.372
40	Deciduous	0.503	0.000
41	Mixed	0.578	0.336
-		RMSE	0.292
-		NRMSE	0.292

Figure 9.39 illustrates the results of the calculations for producing the VE index over the selected sample area. A number of patterns can be observed within the individual compartments. Higher index values are typically present in semi-ancient deciduous woodland. Higher values and greater variability is evident in Frame Wood, whereas more consistent lower values are evident in Tantany Wood, which is an area of higher canopy closure. Plantation and coniferous woodland have typically lower and less variable VE index values than deciduous areas.

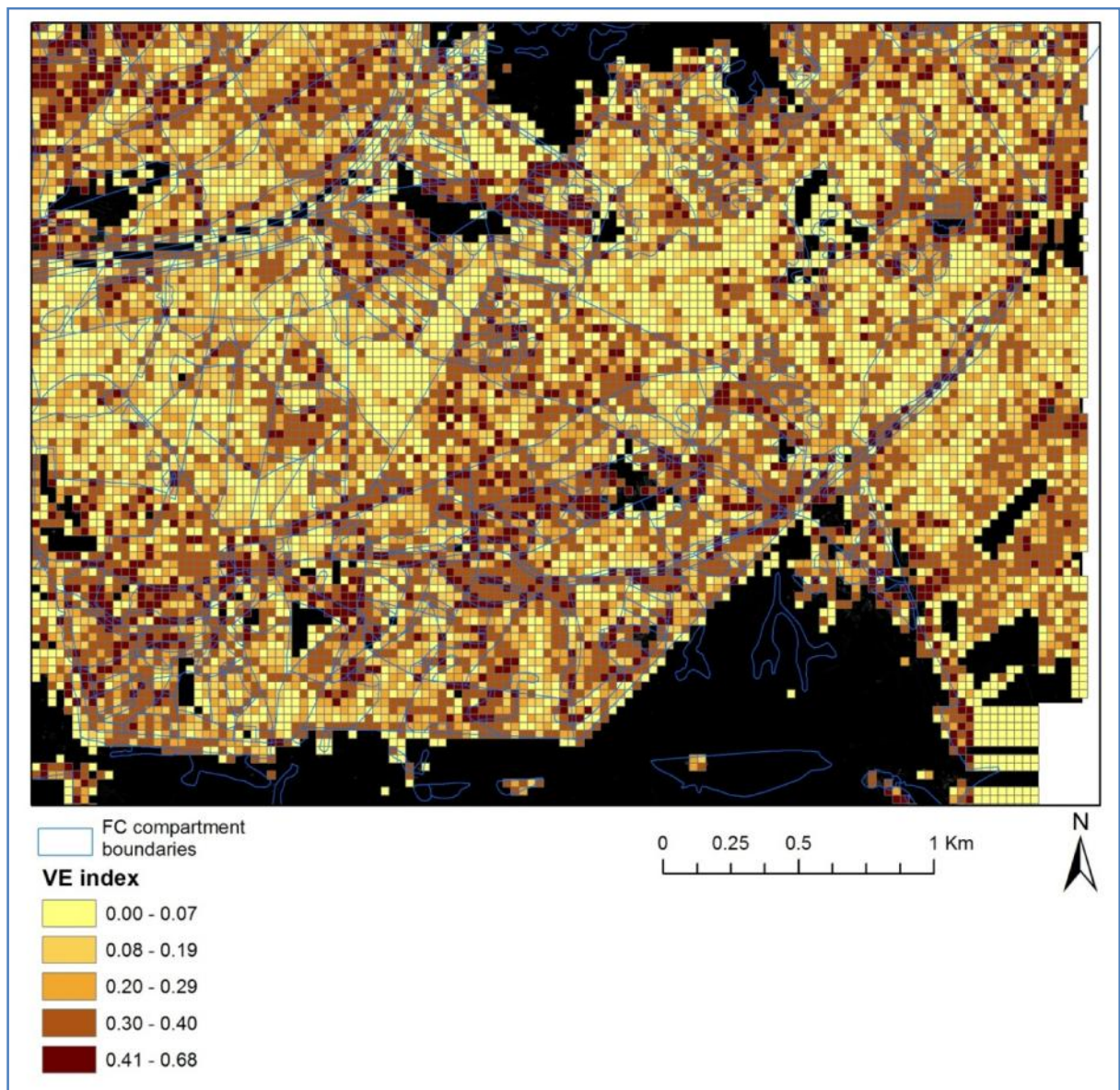


Figure 9.39 – The VE condition index as calculated from ITC information for 30x30m areas.

9.2.2 The Clark-Evans aggregation index

The aggregation index is a measure of the variability in distances between NN tree stems within the field plot extent. As with the VE index above, this index was calculated from the outputs of DR LiDAR ITC analysis data generated from the TIFFS software. Table 9.3 shows the results of the aggregation index calculated for field data and LiDAR derived data. Plots 21 and 31 contained measurement errors and were removed. All of the index results are very close to a value of 1, indicating trees are distributed randomly within the stand (Clark and Evans, 1954; Vorčák et al., 2006). The overall RMSE of the estimated index is 0.236, while NRMSE is 0.254, indicating the predicted values are within 25.4% of the field data equivalent. The largest variation from the field data equivalents are found for plots 24, 25, 26, 27 and 29, which are all deciduous plots, in each of these cases the remote sensing derived index is larger than the field derived index values.

Table 9.3 – Comparison of field vs. remote sensing derived aggregation index

Site no.	Class	Field calc. Aggregation index	DR LiDAR Aggregation index
23	Coniferous	1.338	1.192
24	Deciduous	0.975	1.215
25	Deciduous	0.700	1.168
26	Deciduous	0.516	1.026
27	Deciduous	0.717	1.205
28	Coniferous	1.224	1.237
29	Deciduous	0.828	1.181
30	Coniferous	1.166	1.193
32	Mixed	1.342	1.284
33	Coniferous	1.407	1.279
34	Mixed	1.239	1.328
35	Coniferous	1.166	1.277
36	Coniferous	1.416	1.427
37	Deciduous	1.139	1.277
38	Mixed	1.199	1.199
39	Coniferous	1.377	1.301
40	Deciduous	1.211	1.198
41	Mixed	1.086	1.228
-	-	RMSE	0.236
-	-	NRMSE	0.254

Figure 9.40 illustrates the results of calculating the Clark-Evans aggregation index, as applied to ITC positional information for the 30x30m cells. The index can range from 0 to 2.79 within the map. The values are linked to the horizontal spatial arrangements of tree stems within the 30x30 m cell, where low values indicate a clustered distribution, values around one indicate a random distribution, and higher values indicate a regular distribution. In reference

to Figure 9.38, different forest types exhibit different variability of these index values. Semi-ancient deciduous woodland varies between low and high values, the majority of which fall in the 1.17-1.46 range, indicating a random to regular distribution. Areas in the plantation coniferous woodland are less variable and typically have higher index values in the range of 1.66-1.88, indicative of a more regular distribution. This difference is most obvious when comparing the variability of deciduous areas in Frame Wood to the coniferous area in Hawkhill inclosure to the south.

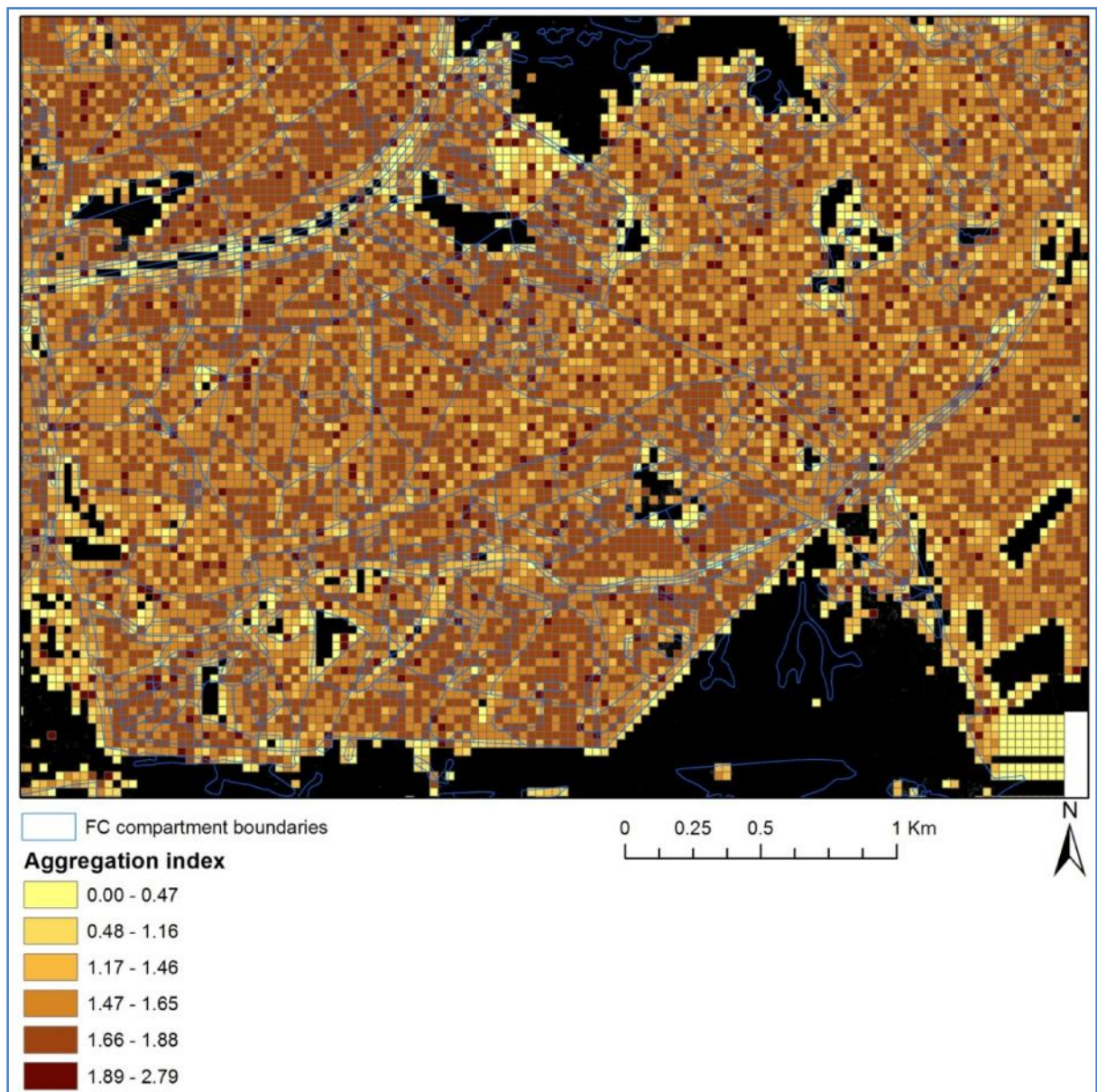


Figure 9.40 – The Clark-Evans aggregation index as calculated from ITC positional information for 30x30m areas.

9.2.3 The tree diameter differentiation index

The tree diameter differentiation index is a measure of the degree of variability in tree stem size within the plot. The inputs for the equation are the number of stems within the plot and an estimate of the sum of the differences between NN tree DBH values, both of which were estimated statistically. The resultant estimation of the index calculations are presented in Table 9.4.

The best remote sensing derived estimate of the index produced an RMSE value of 0.19 and a NRMSE of 0.23. The majority of index estimations are close to the field index value. Plot 14 underestimated the field value, while plot 13 overestimated the field value. Plot 14 was a coniferous plot and dominated by scots pine. There were a very high proportion of tall birch saplings within the plot, which were not taken account of in the index calculation. Plot 13 was a deciduous plot and dominated by oak. This plot contained a number of large standing dead trees which were not taken into account in the field calculated index. The estimate in Plot 12 returned a value of 0; however the field derived index was also very low, indicating the trees within that area were very similar.

The tree diameter differentiation index was calculated for the whole study site, from the statistically derived inputs from the FW LiDAR dataset. The FW dataset proved to best estimate the two components required to calculate this index, see Chapter 8.4. Figure 9.41 illustrates the subset area with the index calculated. The index values range from zero to one, where low values indicate little difference in DBH, whereas values approaching one indicate a large difference in DBH values for the 30x30m plot. There are differences between the values encapsulated within the FC compartments. Those grid-cells within semi-ancient deciduous woodland have a great deal of variation in index values. Larger variation also seem to be present in areas with a more open canopy, such as in Frame Wood. Coniferous woodland generally has very similar values within each compartment, suggesting an even age within compartments, but overall compartments can vary between very low and very high index values.

Table 9.4 – Comparison of the field derived and remote sensing derived tree diameter differentiation index

Site no.	Field calc. index	Best models index
1	0.230	0.315
2	0.192	0.218
3	0.221	0.412
4	0.518	0.560
5	0.264	0.212
6	0.248	0.290
7	0.349	0.339
8	0.362	0.484
9	0.325	0.342
10	0.409	0.216
11	0.454	0.453
12	0.186	0.000
13	0.385	0.752
14	1.000	0.358
15	0.338	0.395
16	0.264	0.329
17	0.232	0.174
18	0.200	0.400
19	0.342	0.308
20	0.401	0.414
-	RMSE	0.189
-	NRMSE	0.232

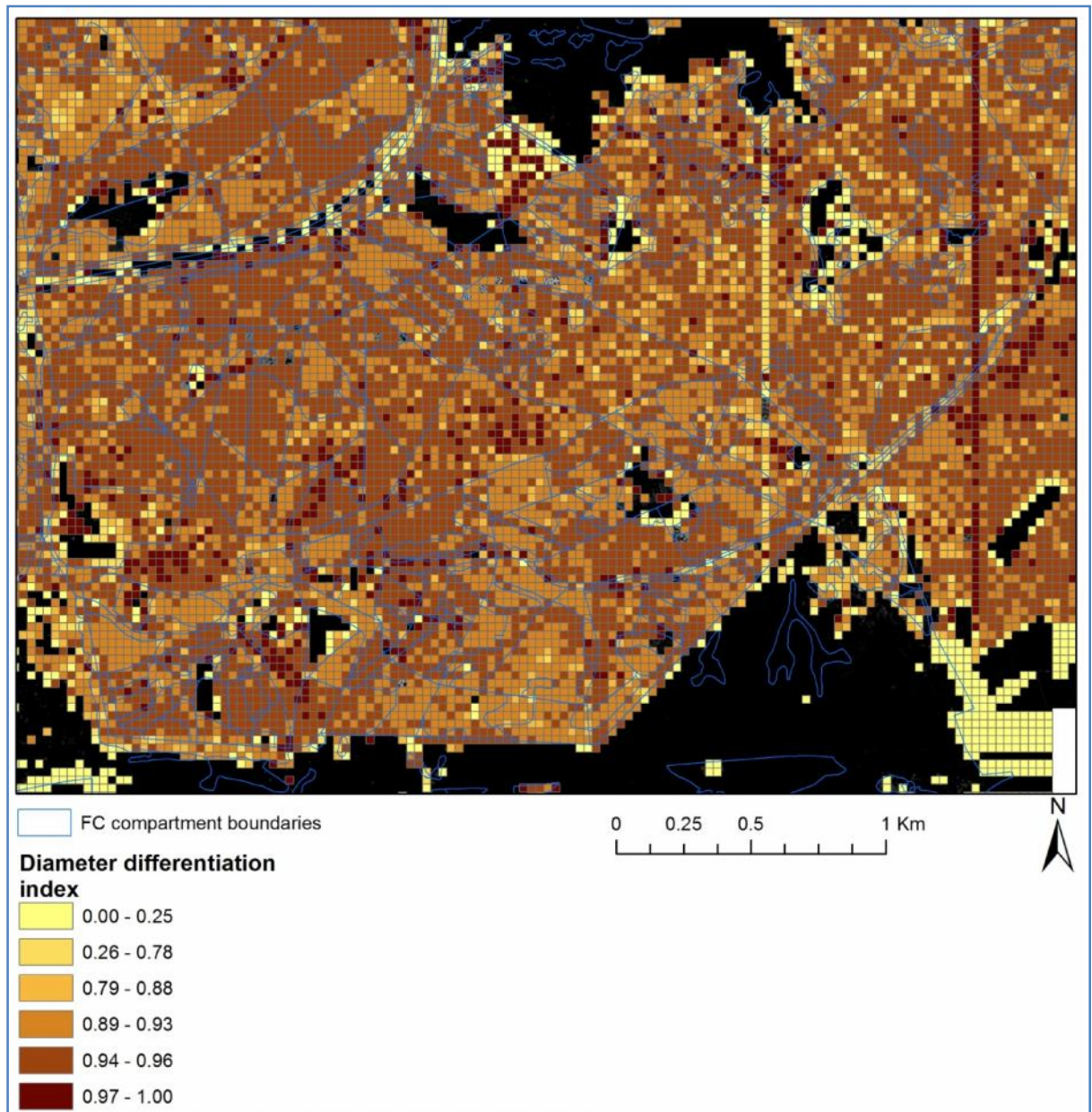


Figure 9.41 – The Diameter Differentiation index as calculated from statistically derived information from FW and hyperspectral datasets for 30x30m areas.

9.2.4 The complexity index (HC)

The Complexity Index (HC) combines four measures of stand description: (i) number of tree species; (ii) stem number per hectare; (iii) dominant height; and (iv) the total basal area per hectare. Each of these inputs was estimated statistically through regression models, as described in the previous step. Table 9.5 summarises the results of the index calculated for the field and remote sensing derived data. There is a great deal of variability in the values for all index calculations, between field and remote sensing derived index values. The remote sensing derived estimate of the HC index produced an RMSE of 113.84 and an NRMSE value of 0.21. The combination of four estimated metrics, each with their own error, could compound this variability; this is also evident in Appendix F section F.2.2. Plot 30 again caused issues and was greatly underestimated, increasing overall model RMSE. The remainder were much better predicted and were within ± 50 of the field index value, i.e. within accuracy 20%.

Table 9.5 – Comparison of field vs. remote sensing derived complexity index

Site no.	Class	Field calc. HC index	'Best' models HC index
22	<i>Deciduous</i>	45.927	29.764
23	<i>Coniferous</i>	67.462	60.028
24	<i>Deciduous</i>	88.474	83.493
25	<i>Deciduous</i>	59.631	108.126
26	<i>Deciduous</i>	64.692	22.132
27	<i>Deciduous</i>	109.911	106.815
28	<i>Coniferous</i>	109.478	42.652
29	<i>Deciduous</i>	38.563	93.241
30*	<i>Coniferous</i>	567.434	111.770
32	<i>Mixed</i>	100.279	45.996
33	<i>Coniferous</i>	137.894	75.419
34	<i>Mixed</i>	106.467	74.156
35	<i>Coniferous</i>	194.753	85.815
36	<i>Coniferous</i>	92.196	78.217
37	<i>Deciduous</i>	45.529	100.459
38	<i>Mixed</i>	120.836	81.291
39	<i>Coniferous</i>	14.841	37.516
40	<i>Deciduous</i>	31.780	71.521
41	<i>Mixed</i>	78.519	75.311
-	-	RMSE	113.842
-	-	NRMSE	0.206

The HC index was calculated for the whole study site, using four statistically derived inputs from the remote sensing datasets. The index is mapped across the subset region in Figure 9.42. The semi-ancient deciduous woodland of Frame Wood received low HC index values. Tantany Wood received similar, albeit more variable values. The highest index values were

found within the plantation coniferous compartment areas, for example within the Frame-Heath inclosure.

The index values range from 0 to 198. A low value of the index would suggest a plot site deficient in many or all of the four inputs, whereas a high value would suggest a plot site with high values for many or all of the four inputs (i.e. species number, stem density, average tree height, or basal area). An example of a low index value would be plot 39, which has one tree species and a relatively low stem density, whereas average tree height and basal area values were high. In contrast an example of a high index value, such as that of plot 35, has higher values for all four inputs.

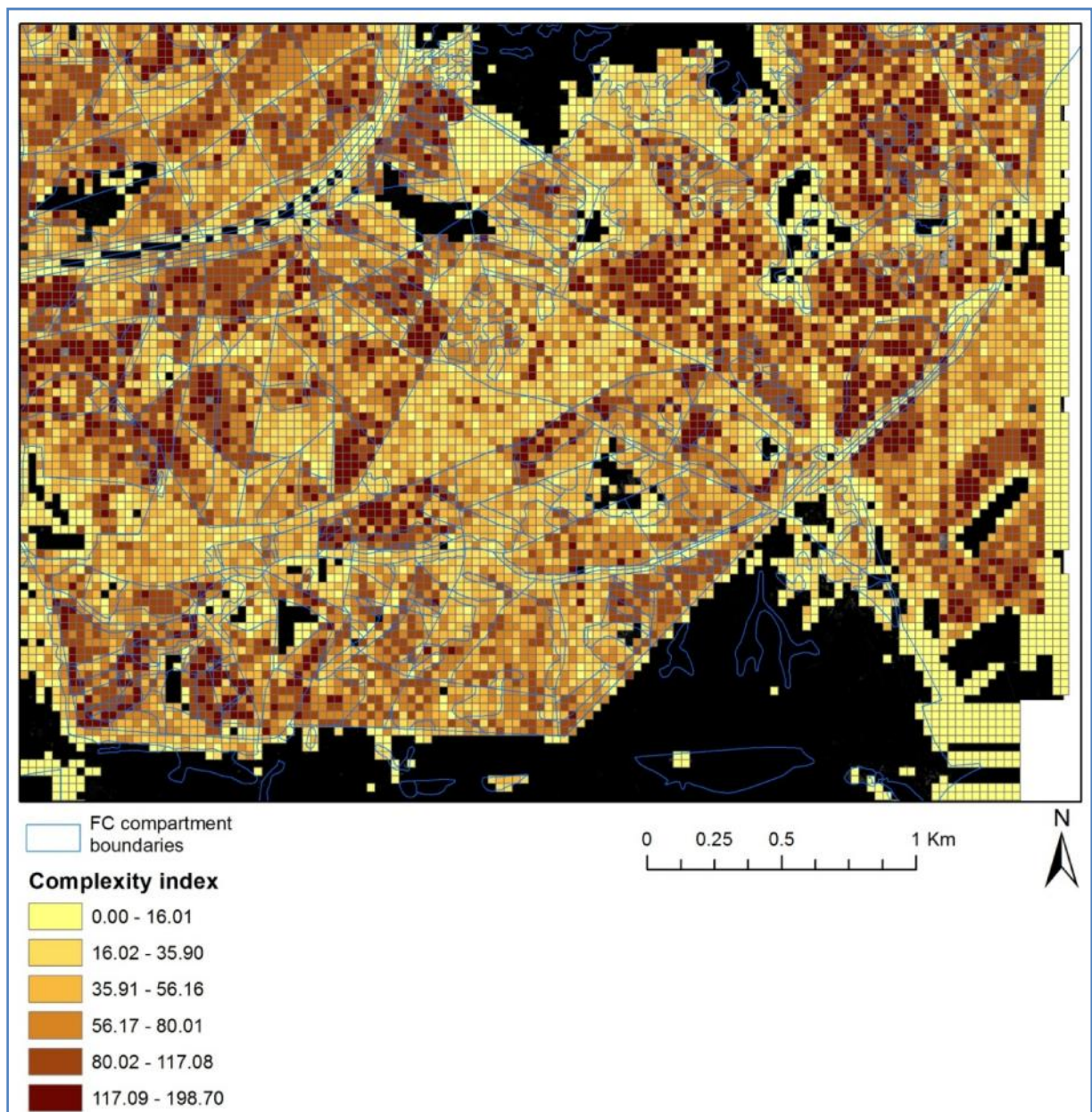


Figure 9.42 – The HC index as calculated from statistically derived information for 30x30m areas.

9.2.5 The complex stand diversity index (CSDI)

The complex stand diversity index (CSDI) also requires a combination of inputs from both direct measurements from remote sensing outputs and statistically derived outputs from regressions. CSDI is composed of four indices relating to species composition, stem size, stem spacing, and crown dimensions. Table 9.6 summarises the results of indices calculated for the field and remote sensing derived data. Plots 21 and 30 contained measurement errors in the field data and were removed as part of the comparison. The remote sensing derived CSDI values produced an RMSE value of 3.59 and an NRMSE value of 0.42

Where there are small field index values (< 8) however, there are overestimations from the remote sensing datasets. For example, field plot 39 has an index value of 5.76, while remote sensing estimates are nearly double (11.30). Conversely, higher field index values are better predicted, such as in plot 34, where the field index value is 12.39 and the remote sensing index value is 11.23. There is variation between remote sensing predictions at higher values; typically this is within ± 3 of the field value.

Table 9.6 – Comparison of field vs. remote sensing derived CSDI index

Site no.	Class	Field calc. CSDI index	Best models CSDI index
22	<i>Deciduous</i>	-	-
23	<i>Coniferous</i>	11.499	10.609
24	<i>Deciduous</i>	12.778	10.127
25	<i>Deciduous</i>	9.920	13.168
26	<i>Deciduous</i>	10.223	10.730
27	<i>Deciduous</i>	11.465	14.045
28	<i>Coniferous</i>	6.882	11.080
29	<i>Deciduous</i>	10.461	11.110
30	<i>Coniferous</i>	-	-
31	<i>Coniferous</i>	4.273	13.077
32	<i>Mixed</i>	9.220	9.680
33	<i>Coniferous</i>	9.505	6.916
34	<i>Mixed</i>	12.390	11.237
35	<i>Coniferous</i>	10.146	8.633
36	<i>Coniferous</i>	7.417	10.259
37	<i>Deciduous</i>	8.376	12.376
38	<i>Mixed</i>	8.779	12.592
39	<i>Coniferous</i>	5.762	11.304
40	<i>Deciduous</i>	8.108	12.078
41	<i>Mixed</i>	10.640	14.620
-	-	RMSE	3.593
-	-	NRMSE	0.422

The CSDI was calculated for the whole study site, using statistically derived and ITC inputs from the remote sensing dataset, and is presented in Figure 9.43. Index values ranged from 0 to 38.58, where higher index values indicate higher heterogeneity in both forest composition and structure within the 30x30 m cell. The areas of semi-ancient deciduous woodlands typically have high values, but also include a lot of variability between high and low values. The highest index values appear to be where there are mixed coniferous and deciduous species in close proximity, or young conifer species. The majority of coniferous plantation areas do not show a great deal of index value variability.

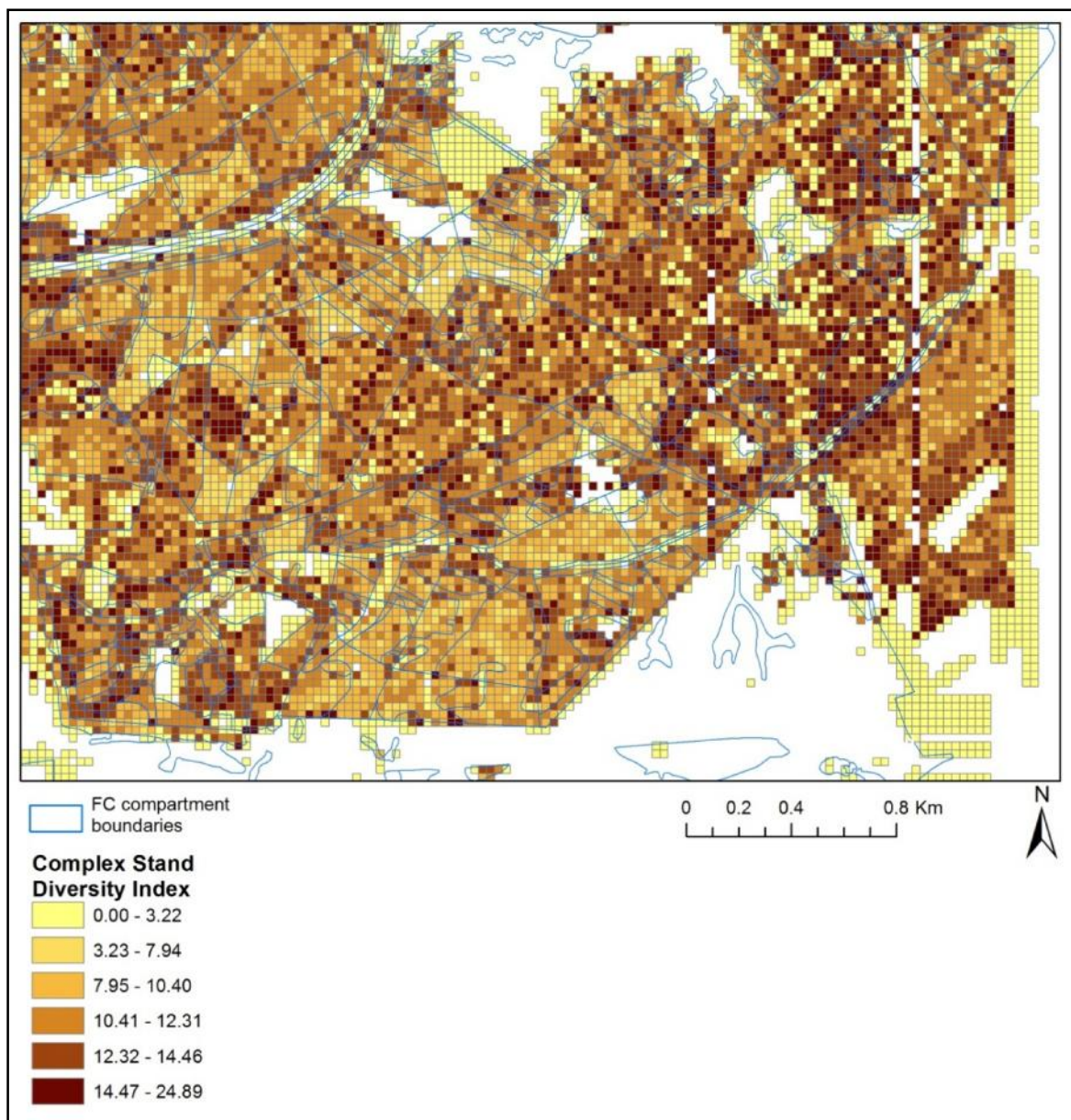


Figure 9.43 – The CSDI as calculated from statistically derived and ITC information from the remote sensing datasets for 30x30m areas.

Given the nature of this combined index, the four component inputs can be broken down and analysed separately. Table 9.7 summarises the differences between the remote sensing and field derived index for the first CSDI component which relates to tree species composition. RMSE and NRMSE values were high, 0.8 and 0.70 respectively, and poorly represented the index value. The index incorporates estimates of tree number, species number and the relative proportions of the most and least abundant species within each 30x30m grid cell, the latter being derived from the number of classified ITC objects (see section 6.4.1). This again highlights the underestimation of tree number from the ITC delineation methods and the statistically derived number of trees.

The construction of the second index element, the index of vertical structure, was derived statistically and reported in section 8.4. The construction of the third index (CSDI-3), the index of spatial distribution, relates to the comparison of the three largest and three smallest NN stem distances, which was reported in section 7.6, Table 9.8 reports the comparison between field and remote sensing derived values of CSDI-3. The inputs determined as the ‘best’ estimates produce RMSE and NRMSE values of 0.07 and 0.45 respectively.

The final index input for the CSDI concerns crown differentiation, the construction of which was assessed in section 7.6 (for ITC derived inputs) and section 7.7 (for estimates of the lowest canopy base height). Table 9.9 reports the relationship between the index values for field and remote sensing derived measures. RMSE and NRMSE values were 1.98 and 0.35 respectively. Overall, remote sensing index values underestimate the field based equivalents by small amounts only.

Table 9.7 – CSDI component “index of tree species composition” – comparison between field and remote sensing derived indices

Site no.	Class	Field calc. CSDI-1 index	Best models CSDI-1 index
22	<i>Deciduous</i>	0.804	0.957
23	<i>Coniferous</i>	0.363	1.239
24	<i>Deciduous</i>	0.318	0.880
25	<i>Deciduous</i>	0.534	1.451
26	<i>Deciduous</i>	0.678	1.218
27	<i>Deciduous</i>	0.513	1.300
28	<i>Coniferous</i>	0.416	1.356
29	<i>Deciduous</i>	0.291	1.386
30	<i>Coniferous</i>	-	-
31	<i>Coniferous</i>	0.000	1.390
32	<i>Mixed</i>	0.783	0.964
33	<i>Coniferous</i>	0.452	0.468
34	<i>Mixed</i>	1.248	1.507
35	<i>Coniferous</i>	0.593	0.843
36	<i>Coniferous</i>	0.494	1.425
37	<i>Deciduous</i>	0.425	1.655
38	<i>Mixed</i>	0.490	1.419
39	<i>Coniferous</i>	0.000	1.191
40	<i>Deciduous</i>	0.463	1.458
41	<i>Mixed</i>	0.660	1.968
-	-	RMSE	0.870
-	-	NRMSE	0.697

Table 9.8 – CSDI component “index spatial distribution” – comparison between field and remote sensing derived indices

Site no.	Class	Field calc. CSDI-3 index	Best models CSDI-3 index
22	<i>Deciduous</i>	-	-
23	<i>Coniferous</i>	0.142	0.112
24	<i>Deciduous</i>	0.134	0.057
25	<i>Deciduous</i>	0.191	0.061
26	<i>Deciduous</i>	0.217	0.062
27	<i>Deciduous</i>	0.219	0.039
28	<i>Coniferous</i>	0.092	0.084
29	<i>Deciduous</i>	0.187	0.063
30	<i>Coniferous</i>	-	-
31	<i>Coniferous</i>	0.063	0.077
32	<i>Mixed</i>	0.067	0.083
33	<i>Coniferous</i>	0.072	0.054
34	<i>Mixed</i>	0.116	0.057
35	<i>Coniferous</i>	0.088	0.055
36	<i>Coniferous</i>	0.069	0.030
37	<i>Deciduous</i>	0.084	0.053
38	<i>Mixed</i>	0.083	0.080
39	<i>Coniferous</i>	0.049	0.061
40	<i>Deciduous</i>	0.080	0.057
41	<i>Mixed</i>	0.134	0.083
-	-	RMSE	0.077
-	-	NRMSE	0.453

Table 9.9 – CSDI component “index of crown differentiation” – comparisons between field and remote sensing derived indices

Site no.	Class	Field calc. CSDI-4 index	Best models CSDI-4 index
22	<i>Deciduous</i>	-	-
23	<i>Coniferous</i>	7.498	3.344
24	<i>Deciduous</i>	8.628	4.438
25	<i>Deciduous</i>	5.049	5.106
26	<i>Deciduous</i>	4.902	3.573
27	<i>Deciduous</i>	6.363	6.573
28	<i>Coniferous</i>	3.555	3.346
29	<i>Deciduous</i>	6.615	3.319
30	<i>Coniferous</i>	-	-
31	<i>Coniferous</i>	2.894	5.195
32	<i>Mixed</i>	3.922	3.527
33	<i>Coniferous</i>	5.650	3.237
34	<i>Mixed</i>	4.900	2.756
35	<i>Coniferous</i>	5.102	3.118
36	<i>Coniferous</i>	3.194	2.361
37	<i>Deciduous</i>	4.327	3.575
38	<i>Mixed</i>	4.431	4.552
39	<i>Coniferous</i>	3.583	4.270
40	<i>Deciduous</i>	3.849	4.059
41	<i>Mixed</i>	5.416	4.271
-	-	RMSE	1.978
-	-	NRMSE	0.345

9.2.6 The target and accumulative scoring technique

The scoring method put forward by Cantarello and Newton (2006) requires the assessment of seventeen compositional and structural metrics against an ‘ideal’ target value. Site metrics were defined through statistical means, the results of which were presented in section 8.4. Table 9.10 summarizes the remote sensing derived index values for all 20 validation sites against the field based equivalents. The index is composed of 17 individual metrics.

The index estimates produced from remote sensing derived metrics produced results with RMSE of 2.5 and NRMSE of 0.25.

Smaller field index values (<7) are overestimated in the case of plot numbers 23, 24 and 38, whereas plot 35 is underestimated. The majority of sites’ field index values were similar to remote sensing derived indices (i.e. within ± 3).

In order to test the individual components of the index it is necessary to view a breakdown of each of the three datasets into the individual scoring, illustrated in Tables 9.11. Plot 24 has the lowest correspondence where only 10 of 17 indicators were correct. Plots 23, 28, 30, 32 and 40 had 11 correspondences. Plots 22, 26, 33 and 38 performed the best with 15-16 out of 17 correct scores. The average correct score was 12.6. In terms of indicators rather than plots, indicator numbers 7, 9, 14, 15 and 17 were 95-100% correct. Indicator numbers 2, 13, and 16 were only 50% correct.

The score index was calculated for the whole study site, using statistically derived inputs from the ‘best’ remote sensing dataset (i.e. those with the lowest RMSE/NRMSE), and is presented in Figure 9.44. For the site-wide calculated index the range was 0 to 15. Semi-ancient deciduous woodland for the whole site received higher index values than coniferous woodlands. Values for Frame-Heath and Tantany Woods varied from 7 to 15. Deciduous woodlands in the Frame-Heath inclosure contain the most consistently high index values. Coniferous plantation woodlands generally contain much less variability, typically 6 to 9.

Table 9.10 – Comparison between the field and remote sensing derived score based index

Site no.	Class	Field calc. score index	Best models score index
22	<i>Deciduous</i>	9	9
23	<i>Coniferous</i>	3	7
24	<i>Deciduous</i>	5	10
25	<i>Deciduous</i>	10	9
26	<i>Deciduous</i>	9	10
27	<i>Deciduous</i>	13	13
28	<i>Coniferous</i>	6	8
29	<i>Deciduous</i>	8	9
30	<i>Coniferous</i>	8	6
31	<i>Coniferous</i>	6	7
32	<i>Mixed</i>	11	7
33	<i>Coniferous</i>	7	5
34	<i>Mixed</i>	10	9
35	<i>Coniferous</i>	10	5
36	<i>Coniferous</i>	8	9
37	<i>Deciduous</i>	7	11
38	<i>Mixed</i>	5	7
39	<i>Coniferous</i>	9	7
40	<i>Deciduous</i>	8	7
41	<i>Mixed</i>	10	11
-	-	RMSE	2.510
-	-	NRMSE	0.251

Table 9.11 – Decomposition of the combined ‘best’ dataset score index

*Differences between estimated and field recorded are indicated in yellow.

Indicator	site 22	site 23	site 24	site 25	site 26	site 27	site 28	site 29	site 30	site 31	site 32	site 33	site 34	site 35	site 36	site 37	site 38	site 39	site 40	site 41
1. No. of trees (ha ⁻¹)	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1
2. Shannon index for native trees (i)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1
3. Basal Area (m ² ha ⁻¹)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1
4. Mean diameter (cm)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1
5. Standard dev of diameters (cm)	1	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	1
6. Percentage of big trees	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7. Mean height (m)	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8. No. of total saplings (ha ⁻¹)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9. No. of native saplings (ha ⁻¹)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10. Vol. downed dead wood (m ³ ha ⁻¹)	1	1	1	1	1	1	1	1	0	0	1	0	1	0	1	1	1	0	0	1
11. Downed dead wood decay class	1	1	1	1	1	1	1	1	0	0	1	0	1	0	1	1	1	0	0	1
12. Volume of snag (m ³ ha ⁻¹)	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13. Snag Decay class	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14. No. of native seedlings (ha ⁻¹)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15. No. of native seedlings (ha ⁻¹)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16. Shannon index for native seedlings (i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17. No. of ground veg. species	9	7	10	9	10	13	8	9	6	7	7	5	9	5	9	11	7	7	7	11

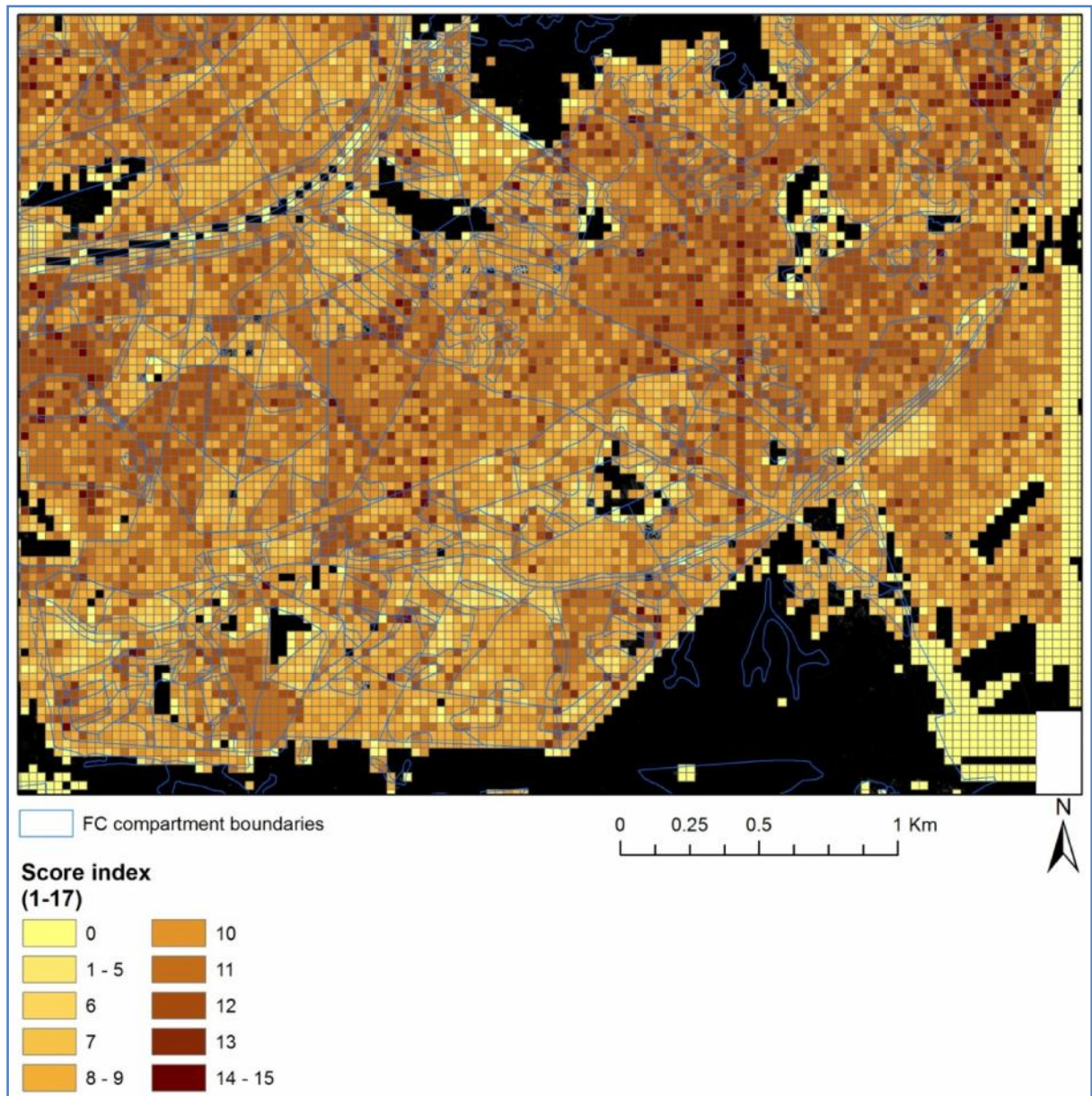


Figure 9.44 – The score index as calculated from statistically derived information from FW and hyperspectral datasets for 30x30m areas.

9.2.7 Condition index overview

Both the VE and aggregation index were derived from ITC metrics, but contained a number of flaws due to the underestimation of stem number and the non-detection of suppressed trees. Until these flaws in the individual tree delineation process can be rectified the approach to derive these indices may be in error.

The diameter differentiation index utilises the fewest inputs in comparison to the other three indices and overall patterns can be discerned visually. Much of the map is however very similar in cell value, whereas there is a great deal of variability in semi-ancient deciduous woodland areas. While this is entirely feasible, the distinction between different forest types is often problematic.

The HC index utilises the second fewest inputs in comparison to the other indices here, and shows clearer distinctions between compartments and forest structural types. It does only seem to distinguish between low and high values however. In addition it has been demonstrated that high field index values are underestimated by the remote sensing index values.

The CSDI index is the most computationally complex of the condition indices, and incorporated a number of individual tree based metrics. The standardised map shows a mix of distinct groupings, typically in coniferous woodlands, and variability, especially in deciduous areas. There are additional concerns over the use of a number of ITC based metrics, in particular those required for calculating the relative species proportions.

The score-based index utilises the most input metrics, 17 in total. Similarly to the complexity index, there are more distinct and even groupings throughout the study site. Overall remote sensing derived estimates correspond well with field data calculated values, producing an RMSE of 2.5. The mapped index values corresponded well to the different FC compartment boundaries and structural types. Its use of area-based metrics instead of direct ITC metrics reduced a number of uncertainties. Bearing in mind sources of error and its ability to discriminate between the different forest structural types when considering map products.

9.3 A cross-comparison of condition index results

The following sections contain the results of a standardisation method applied to all calculated condition indices and a comparison of the resultant maps. Five categories were assigned to the mapped data, which corresponded to 1 through 5 for low to high index values (6 categories including zero). The index values and range of values are different for each index, thus each of the following sections includes a key identifying which values conform to which category for each condition index value. It should be noted that equal intervals over the

range of the index were used for each of the categories. For each of the following steps, the operations were carried out using both the R and ArcMap software.

9.3.1 Index categorisation

For the sample area shown in Figures 9.45-9.50, the percentage in each of the five categories of condition index values is given in Table 9.12. It should be noted that there were two obvious errors present in each of the maps produced for this section which could not be rectified. Both of these errors lie to the east of the subset area and run north to south, and were caused by errors in extracting and mosaicing FW LiDAR metrics. For the two condition indices derived through ITC-metrics, the VE and aggregation indices, the majority of grid-cell values occupy categories 1 and 2, with very few exceed category 3. In interrogating the VE index map (see Figure 9.45) spatial patterns can be observed, which conform to compartment boundaries; however the map of the aggregation index (see Figure 9.46) does not show any spatial pattern. The diameter differentiation index map has the majority (84%) of grid-cell values occupying the 5th category; again the mapped values (see Figure 9.47) do not show any spatial pattern.

The remaining three categorised condition indices (see Table 9.12) have a spread of values across many of the five categories, and definite spatial patterns, within compartment boundaries, can be observed in the mapped data. The HC index (see Figure 9.48) has a range of values present in both semi-ancient and plantation woodland compartment types. For example Frame Wood contains many category 1 grid-cell values whereas Tantany Wood contains many category 4 and 5 values, in addition to a number of category 1 values. Hawkhill inclosure likewise contains variation between the different coniferous compartments featuring the full range of values; however grid-cell values are typically consistent within compartments. The CSDI (see Figure 9.49) has a great deal of variability in index categories present in the semi-ancient deciduous woodland which also conform to compartment boundaries. CSDI category grid-cell values located within plantation inclosures contain much less variability. The score-based index map (see Figure 9.50) shows more consistent category values, even for semi-ancient deciduous woodlands. The score-based index typically has higher values for semi-ancient deciduous woodlands than for coniferous woodland.

Table 9.12 – The percentage of each of the categories for the six condition indices

Index category	VE index (%)	Agg. Index (%)	Diam. Diff. index (%)	HC index (%)	CSDI (%)	Score-based index (%)
No value	25.97	5.43	12.09	12.34	18.52	12.00
1	17.08	20.61	0.08	17.16	0.44	003
2	45.83	64.82	0.21	29.10	1163	9.37
3	1107	7.97	0.46	19.71	34.20	48.22
4	0.04	1.15	2.68	10.56	23.80	30.35
5	0.00	0.02	84.47	11.14	11.40	0.02

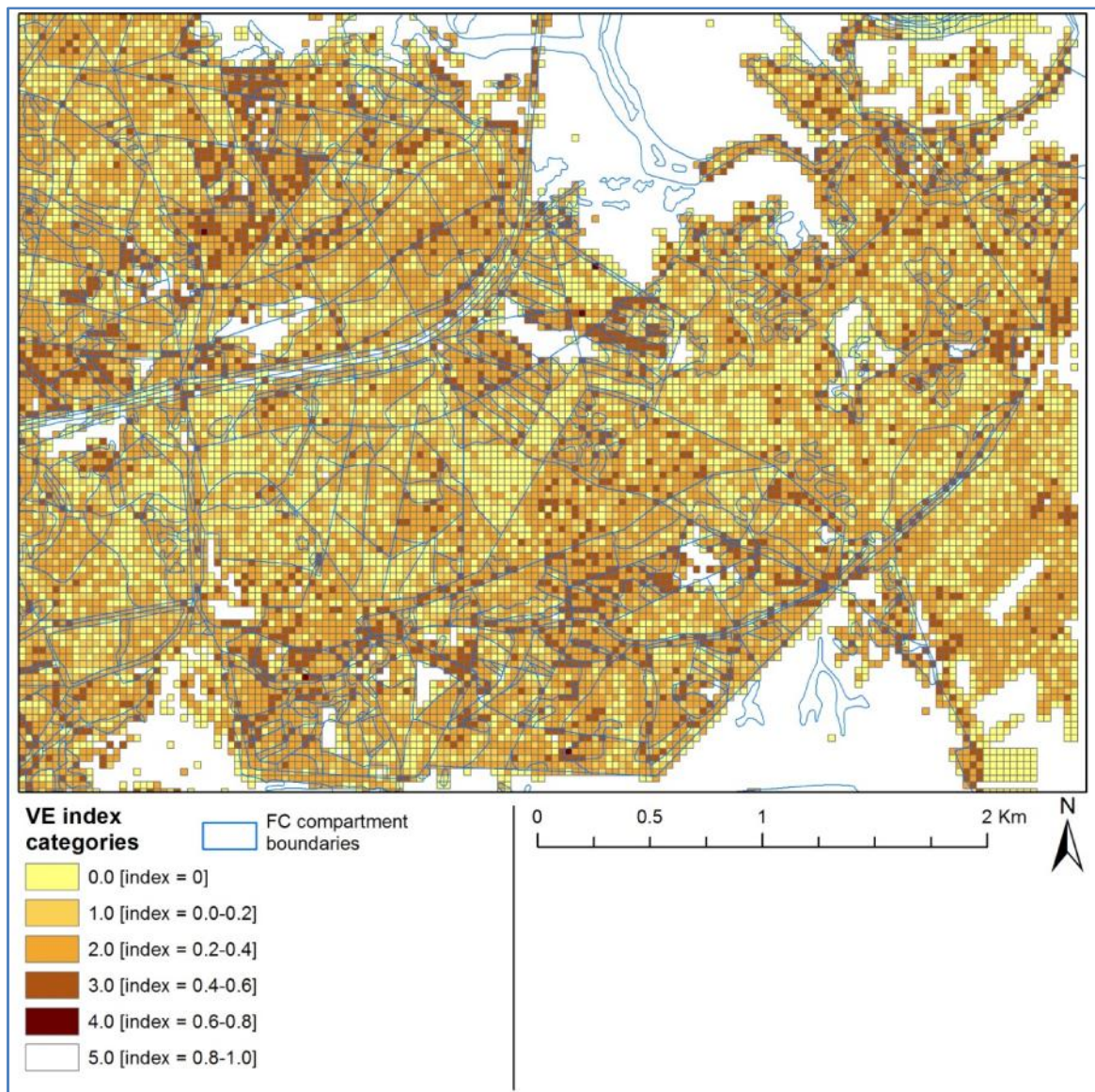


Figure 9.45 – A subset of the study site depicting the categorised VE index results.

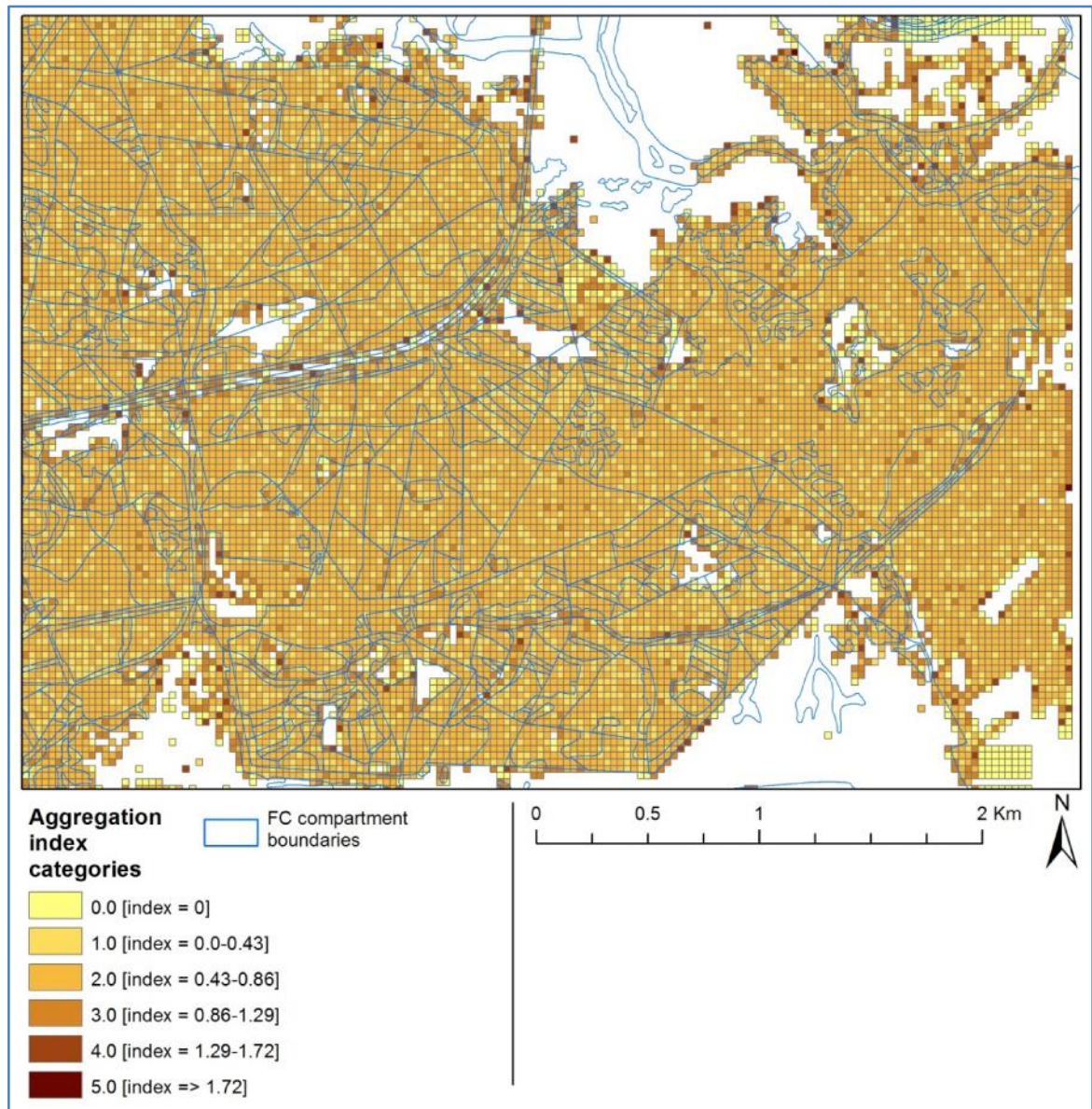


Figure 9.46 – A subset of the study site depicting the categorised Clark-Evens aggregation index results.

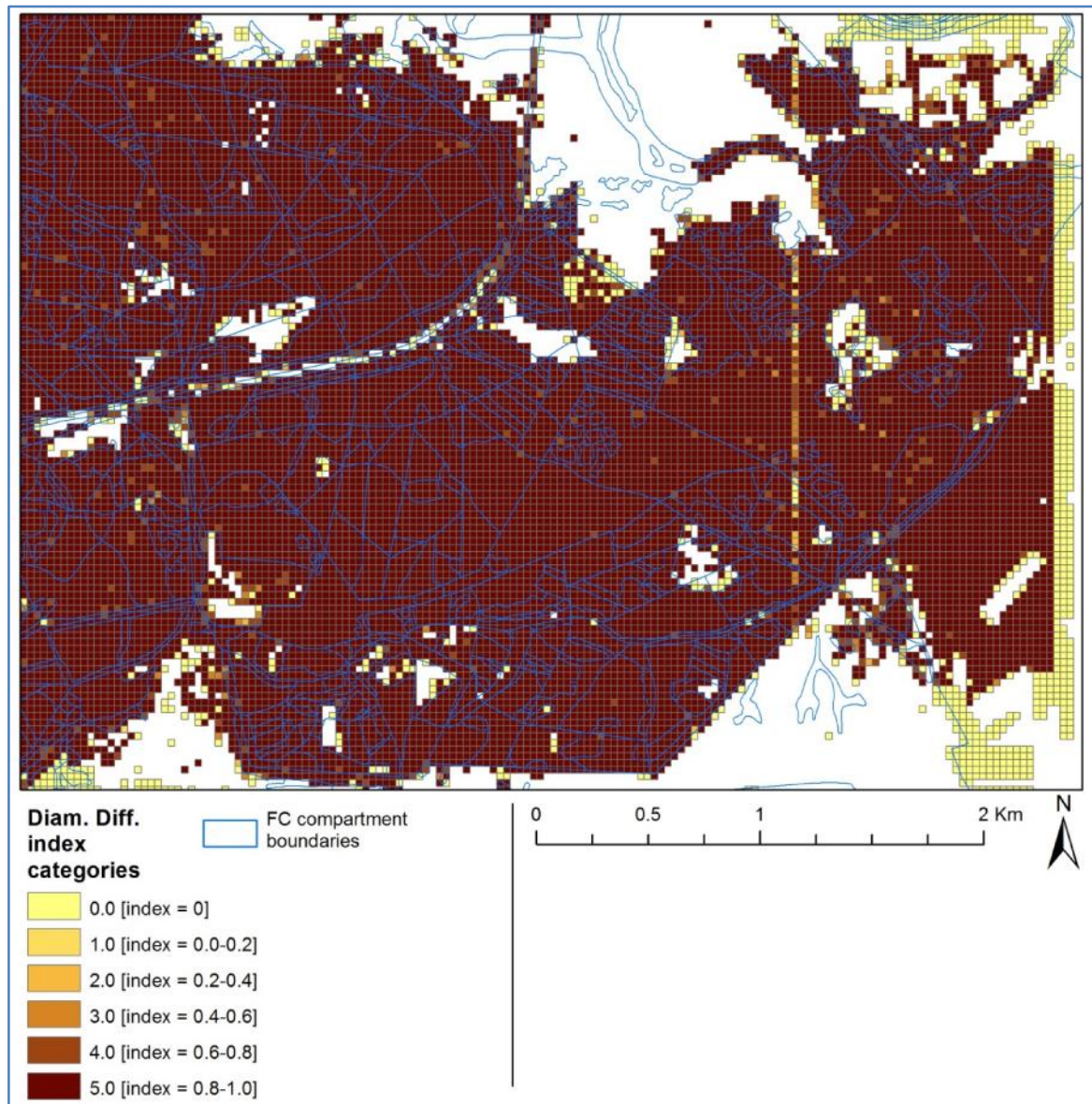


Figure 9.47 – A subset of the study site depicting the categorised diameter differentiation index results.

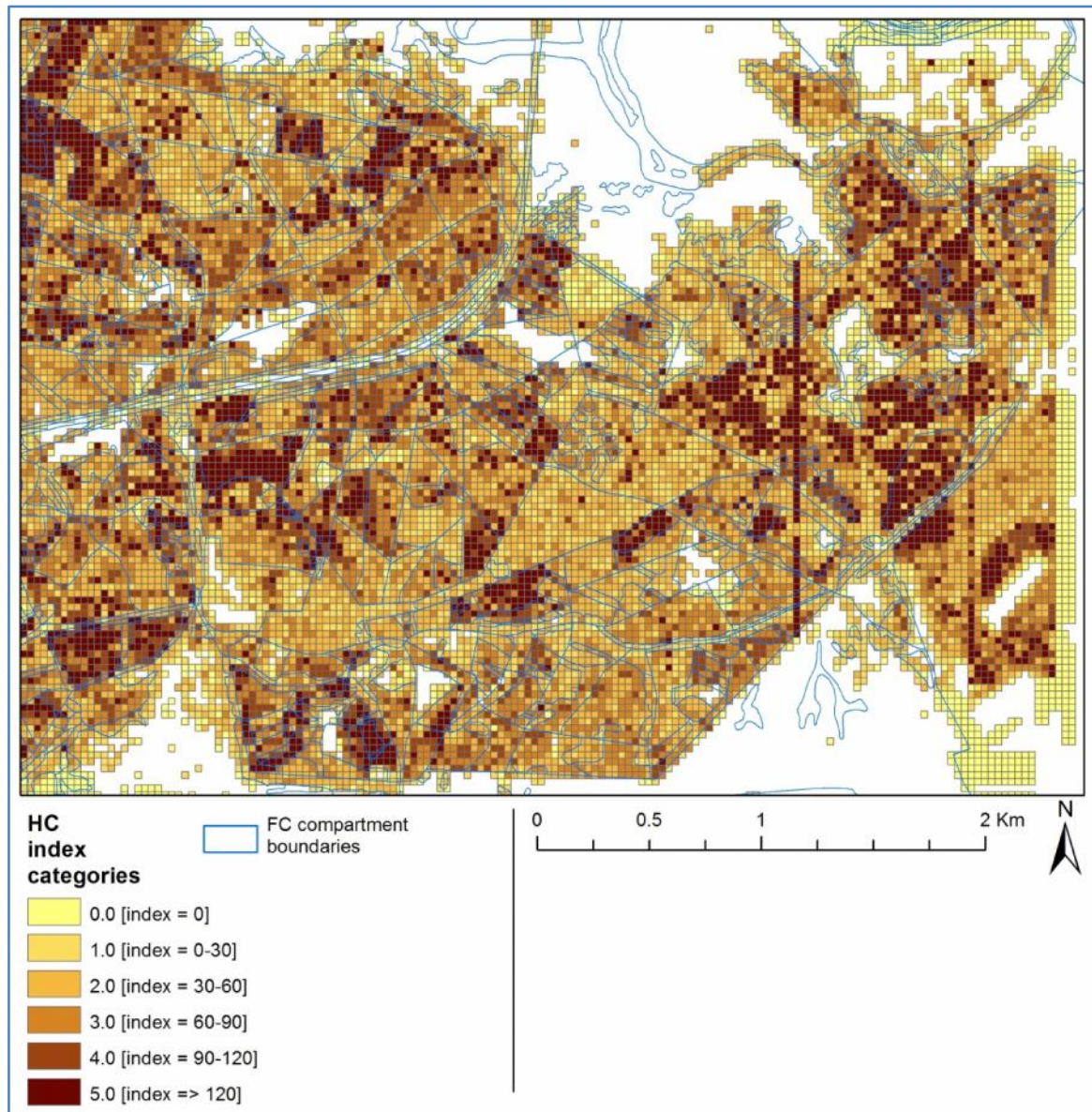


Figure 9.48 – A subset of the study site depicting the categorised complexity index (HC) results.

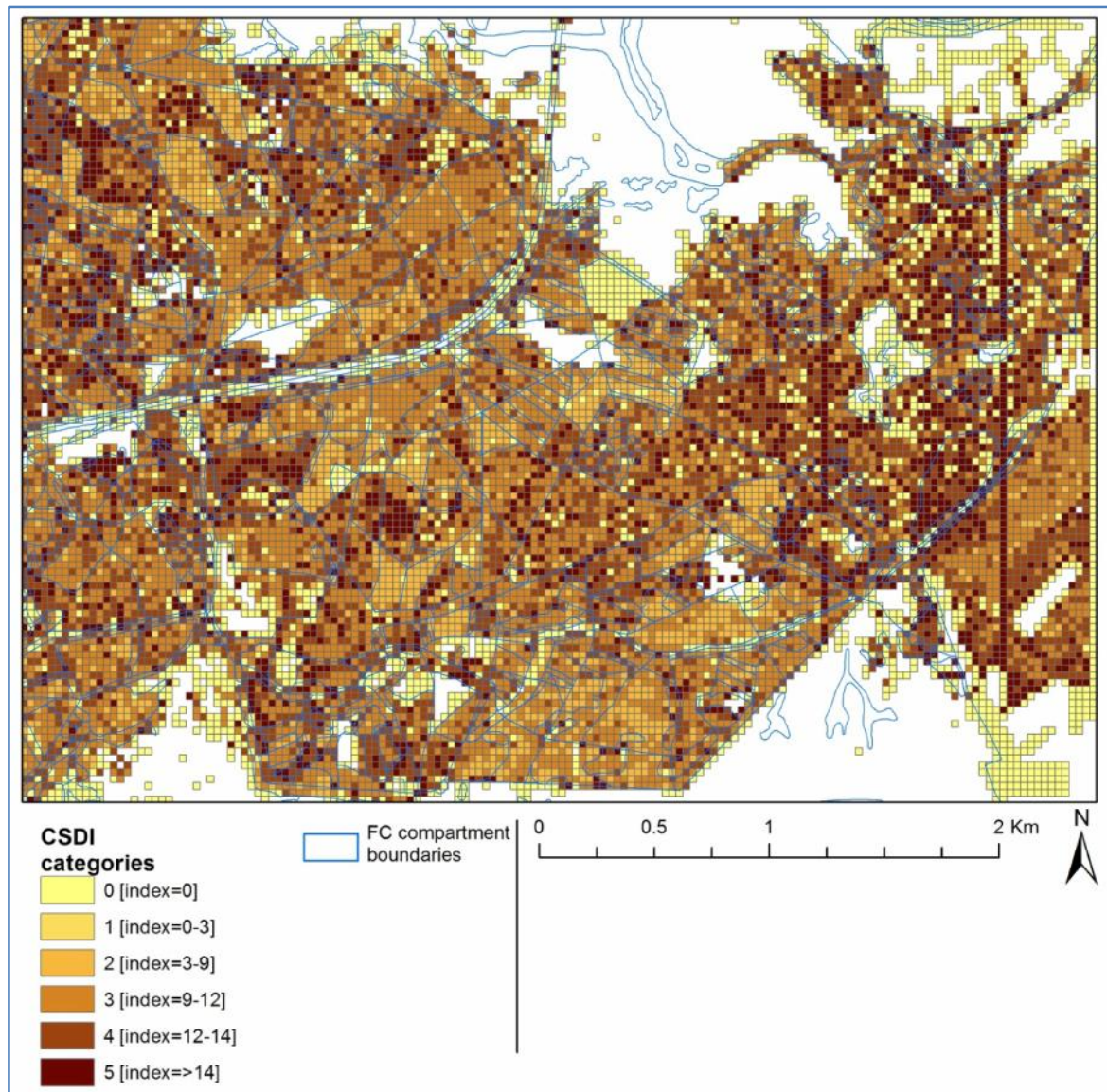


Figure 9.49 – A subset of the study site depicting the categorised CSDI results.

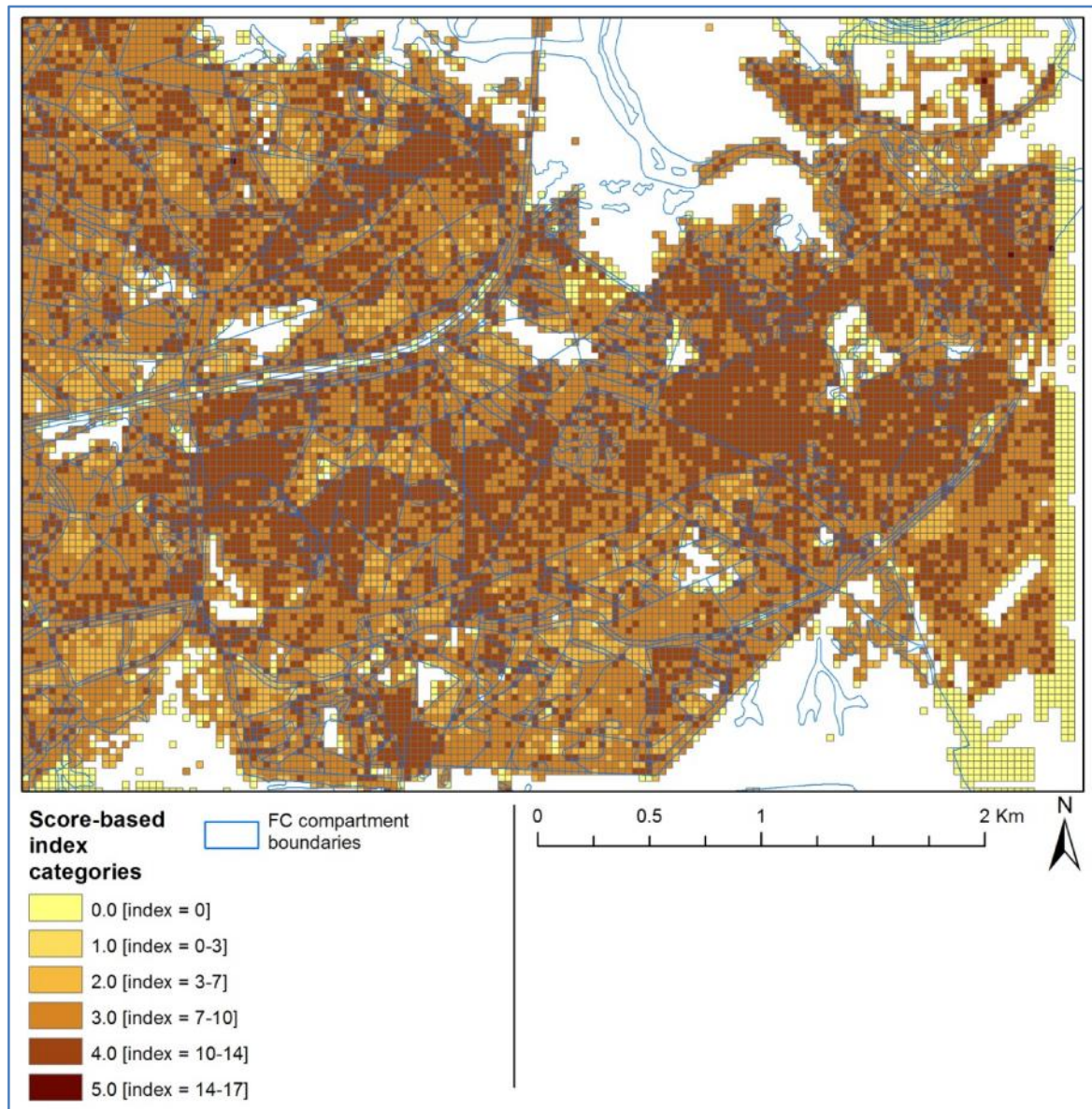


Figure 9.50 – A subset of the study site depicting the categorised score-based index results.

9.3.2 Assessment of index mapping results

The overall patterns of index values were preserved through the categorisation. As highlighted in the previous section, the VE index may not provide trustworthy results in areas of closed canopy, for example in many of the semi-ancient deciduous woodlands. The Clark-Evans aggregation index produced many similar values, only providing a general distinction between coniferous and deciduous woodlands. The categorising of the results by equal intervals over the range of the index illustrated that the distinction between woodland types was very small. The range of values for the diameter difference index was also shown to be very small; so much so that no features could be discerned when categorisation was

implemented. Due to the issues encountered with the VE, Clark-Evans aggregation, and diameter difference, these indices were removed from the subsequent steps.

The conditional maps for the HC, CSDI and the score-based indices show distinct patterns, and include variability of index values within semi-ancient and plantation woodlands. In order to compare the mapped distributions of standardised index values two calculations were applied combining all three index map layers, providing the mean and the standard deviation of index values for each grid-cell. The ultimate aim of this assessment was to test if high values in one mapped index layer were consistent with the others and judge if the indices were showing similar or differing information.

The mean of each of the three categorised indices was computed for each 30x30m grid cell, the result of which is illustrated in Figure 9.51. For the majority of deciduous woodlands, high values are present in their intersecting grid-cells. Frame Wood possesses lower values however, as in the HC index. Coniferous area mean values are typically lower than for deciduous areas, a typical pattern exhibited for all three indices. In addition coniferous areas typically show less variable values within each compartment extent than for deciduous areas.

A total of six sample areas are annotated in Figure 9.52, labelled 1-6, for three areas where high mean values and three areas where low mean values are encountered.

1. The first of the annotated areas is located within Tantany Wood, where all three index results are high values. This sample area is an oak and beech dominated semi-ancient woodland (planting date ~1800). There is also a great deal of variability in values encountered within Tantany Wood, which range from low (0) to high (5). The four HC index components of stem density, basal area, canopy height and species number are all high. The four components of the CSDI are all high for the variability of DBH, stem NN space, and crown size and species proportions. The score-based index utilises some of the tree size and variability metrics covered in the previous two indices and indicates sufficiently high (i.e. beyond the metric target thresholds) for 10-11 metrics (including deadwood volume and decay class), whereas understorey metrics (e. g. seedling number) were too low to meet the threshold.

2. The second annotated area of high mean values is located within the west of the Frame-Heath inclosure. This woodland is dominated by beech and oak woodland of different planting dates, 1938 and 1852 respectively. Of the four HC index components species number and height are high, whereas the stem density and basal area metrics range from moderate to high. The CSDI index components for species and crown size diversity are high, whereas DBH and NN space variability is medium to high. For the score-based index, as with the previous subset in Tantany Wood, 11 targets are achieved relating to tree size, variability and composition, whereas understorey metrics are typically below their respective targets.
3. The third annotated area containing high mean values is located within the Frame-Heath inclosure. This compartment is a mix of scots pine, norway spruce and birch species planted between 1940 and 1960. It should be noted that the mean values within this compartment were slightly lower than for the deciduous subset areas. The HC index components of species number and stem density were all high, whereas canopy height and basal area was moderate. For the CSDI, species diversity was high, crown diversity was moderate, DBH differentiation was low to moderate, and tree NN spacing was low. For the score-based index the majority of targets were achieved relating to tree size, variability and composition. Standing deadwood volume and many metrics related to understorey, failed to achieve the scoring targets.
4. The fourth annotated area has low mean values. The area is located in Denny-Lodge inclosure. The compartment is dominated by corsican pine trees which were planted in 1968. The HC index components are typically low to moderate values. The species number is very low, either 1 or two species. Stem density, basal area and canopy height are moderate values. The three CSDI components for species diversity, DBH and NN spacing differentiation are low, and the final component, crown differentiation is only moderate. Of the 17 inputs to the score-based index, only 6-7 metric targets are achieved. Where stem density, basal area, average DBH and the standard deviation of DBH vary in-between neighbouring cells of being just above or below the target thresholds. The percentage big tree, standing decay and downed deadwood decay class metrics are consistently scored highly.

5. The fifth annotated area which has low mean index values is located in the Frame-Heath inclosure. This compartment was dominated by scots pine and corsican pine coniferous species which were planted in 1944. The HC index values are low for this compartment, where each of the four input components was low to moderate. Likewise the input components for the CSDI were low to moderate, producing low index values overall. Many of the 17 inputs for the score-based index which related to stem density and tree size were very close to the target values, however species composition, deadwood volume, understorey proportion and composition metrics were very low or non-existent within each 30x30m cell and thus were scored 0. This culminated in a low index value.
6. The sixth annotated areas which had a low mean index value was located in Frame-Heath inclosure. The two compartments were dominated by scots pine, with small proportions of corsican pine which was planted in 1944. The HC index was low, where the components were low to moderate in value. The CSDI also produced low values. The CSDI components were all low, indicating little diversity. The score-based index was also low. Where again some of the input metrics relating to stem density and tree size were very close to the target values, however species composition was again low, as were deadwood volume and understorey metrics.

There is a high level of correspondence between the three condition index assessment techniques in terms of areas with high and low values. The highest index values are almost exclusively found in deciduous dominated areas, with only a limited number of coniferous dominated compartments exhibiting high values. Older deciduous woodland dominated by beech and oak species, even when mixed with newer planted deciduous trees produced the highest results consistently. The presence of large trees (i.e. DBH values) with variable sizes, with higher number of species and the presence of deadwood, define these areas. These areas may also contain understorey components lacking in coniferous areas. The highest values in coniferous areas were for compartments with a mix of species types. These areas were also defined by the presence of large trees with variable sizes, higher numbers of species and the presence of deadwood. Table 9.13 lists each of the compartments for which the majority of the grid-cells intersecting it had mean values of greater than four (for the subset area in Figure 9.52), species type, percentage cover and planting year from the FC compartment database are also given. A total of 10 compartments fulfilled this criterion, the majority of

which were a combination of oak and beech tree species with long continuity, where planting dates begin in 1800. There was only one small coniferous compartment which fulfilled this criterion, which consisted of a mix of two conifer species (scots pine and norway spruce) and one deciduous species (birch).

The majority of coniferous areas exhibited relatively small tree sizes, often with little variation in their respective sizes, few species with little or no deadwood or understorey. Table 9.14 lists each of the compartments for which the majority of the grid-cells intersecting it had mean values of less than two (for the subset area in Figure 9.52); again species type, percentage cover and planting year from the FC compartment database are given. For the ten identified compartments, the majority of these compartments contained a high proportion of corsican pine species from recent planting dates. Each of the compartments indicated with no detail of percentage cover or planting year in the FC database were felled shortly after the acquisition of the airborne data.

Table 9.13 – Compartment composition for high mean values (> 4) within forest subset (see Figure 9.51)

	Species	Percentage cover (%)	Planting year
1	oak/beech/scots pine	45/29/27	1861/1861/1928
2	oak/beech	50/50	1829/1956
3	Mixed broadleaved	100	1861
4	beech/oak	51/49	1938/1852
5	beech/oak	53/49	1938/1852
6	oak/beech/birch	57/33/10	1852/1940/1940
7	scots pine/norway spruce/birch	35/30/21	1940/1960/1960
8	oak	100	1852
9	oak/sweet chestnut	90/10	1809/1809
10	Mixed broadleaved	100	1961
11	oak/beech	55/45	1800/1800
12	beech/oak	61/39	1800/1800

Table 9.14 – Compartment composition for low mean values (≤ 2) within forest subset (see Figure 9.51) ("?" = not stored within the FC database)

	Species	Percentage cover	Planting year
1	corsican pine	100	1968
2	corsican pine	100	1968
3	corsican pine	100	1968
4	corsican pine/scots pine	?/? ($\Sigma = 100$)	?/?
5	Mixed broadleaved/mixed conifers	80/20	2001
6	corsican pine	100	1971
7	scots pine/corsican pine	?/? ($\Sigma = 100$)	?/?
8	scots pine/corsican pine	?/? ($\Sigma = 100$)	?/?
9	scots pine/corsican pine	71/29	1948/1948
10	corsican pine/stika spruce	56/24	1982/1986

The primary driving factors it would seem are related to the average size of the trees (i.e. DBH) and the variability of the sizes over the plot which is especially present in oak and beech woodland with long continuity, for example with planning dates approaching 1800. Given the correct conditions much younger coniferous compartments can reach similar values. The presence and volume of deadwood seems to follow these patterns. Understorey metrics such as vascular species number, the number and composition of seedlings and saplings does not seem to follow however. Due to the limited levels of understorey encountered within the New Forest study site, finding a link may not be possible.

The second calculated map depicts the standard deviation of each of the grid-cells, see Figure 9.53. This map highlights the inconsistencies between the three derived index maps. The standard deviations range from 0 to 2.5, where the majority of standard deviations exist between the range of 0.5 and 1.5. The largest deviations are present within the semi-ancient woodlands and the deciduous areas within the plantation inclosures, with the largest areas existing within Tantany Wood and Frame-Heath inclosure. High standard deviation values exist for coniferous woodland compartments as well, an example of which is located in the New-Copse inclosure.

Five example areas with high standard deviation values are annotated in Figure 9.54. The areas are labelled 1-5, the initial four of which are deciduous woodland areas and the remaining one is a coniferous area.

1. The first of the annotated areas is located within Tantany Wood. This sample area is an oak and beech dominated semi-ancient woodland (planting date ~1800). The HC and CSDI index both exhibit variability within this region of both high and low values, whereas the score-based index values are consistently high. The CSDI is sensitive to areas where large canopy gaps occur or low canopy heights, which correspond to smaller index values. Likewise there are areas of low HC values which follow a similar pattern. While score-based index values are higher, it does exhibit lower values in the same areas, which indicates the score-based index is dependent on other attributes.

2. The second deciduous area is located within Frame Wood. The inclosure is dominated by oak, with a small proportion of beech, with plantation dates between 1750 and 1800. Here the HC index is consistently low, which is uncommon for deciduous woodland. The CSDI and score-based indices however have higher, albeit variable, index values for this woodland type. The HC index is based primarily upon forest parameters of size, stem density, canopy height and basal area, in addition to number of tree species. Stem density and basal area are relatively low in comparison to other areas, such as Tantany Wood. The CSDI is based upon the diversity of structural parameters; these are DBH size, differences in stem spacing, and crown size difference. Each of the CSDI components are relatively high in this region. The score based index value is higher due to the high presence of deadwood and occasionally sufficient numbers of saplings to go above the target value.
3. The third deciduous area is located within the north of the Denny-Lodge inclosure. The compartment is dominated by oak, with approximately 26% beech, planted in 1870. As with the second sample area the HC index is consistently low, and the CSDI and score-based indices are higher. The HC index components for stem density were very low, which would account the low index value. CSDI inputs for structural diversity (DBH, stem spacing and crown size) were high. The score-based index inputs exceeded the target values for many of its 17 inputs, which included high levels of deadwood and intermittent high concentrations of saplings.
4. The fourth deciduous area is located within the Denny-Lodge inclosure and is a mix of broadleaved tree species surrounded by coniferous plantations. Plantation dates vary from 1852 to 1928, even within the same compartment. As before there is an obvious difference between the three index values, the HC and score-based values are high, whereas the CSDI value is lower. For example the HC components based on size parameter (tree density, height and basal area) were all high, whereas the CSDI inputs for structural diversity (DBH, stem spacing and crown size) varied between low, moderate and high. The areas of largest variation are coincident with large canopy gaps.

5. The fifth area is located in the New-Copse inclosure which is a coniferous dominated compartment. The compartment contains western hemlock primarily, various conifers and a number of sweet chestnut species, the planting date was 1965. Both the CSDI and score-based index produced moderate values for this compartment, whereas the HC index produced high values. The HC index inputs all range from moderate to high values resulting in a high index value. The CSDI inputs are typically low for this compartment extent. While the score-based index shares some of the same inputs as the HC index, their significance within the index calculation is not as high, and as such these inputs and 13 others, result in only a moderate index value.

There is a great deal of similarity between the three indices; it is interesting to note that deciduous woodlands typically receive higher values than coniferous woodlands. It should be noted that few grid-cells have a very high (5) index value for the score-based index, which is responsible for the many of the small differences present in the standard deviation map. It is clear, even though the differences are small, that the three condition indices (HC, CSDI and score-based) reflect different elements of the composition and structure of the woodlands in question, especially when applied to deciduous woodland with long continuity (i.e. planting dates of ~1800). The HC index is more sensitive to tree size parameters, where it tends towards an almost binary output depicting coniferous areas and deciduous woodland types, as low and high respectively. The CSDI on the other hand relates more to structural diversity which can transcend either the broad types of deciduous or coniferous woodland, but tends to higher and more variable values in deciduous areas. A major concern is that the CSDI overestimates or underestimates the index value when large canopy gaps (i.e. where areas of ground are visible in the CHM) are encountered within grid-cells and may be a result of using ITC metrics directly. This issue was present in small areas however. The score-based index combined both tree size and diversity metrics at a relatively basic level in addition to many other factors. The score-based index also did not directly use ITC-metrics and thus should be less influenced by its associated error.

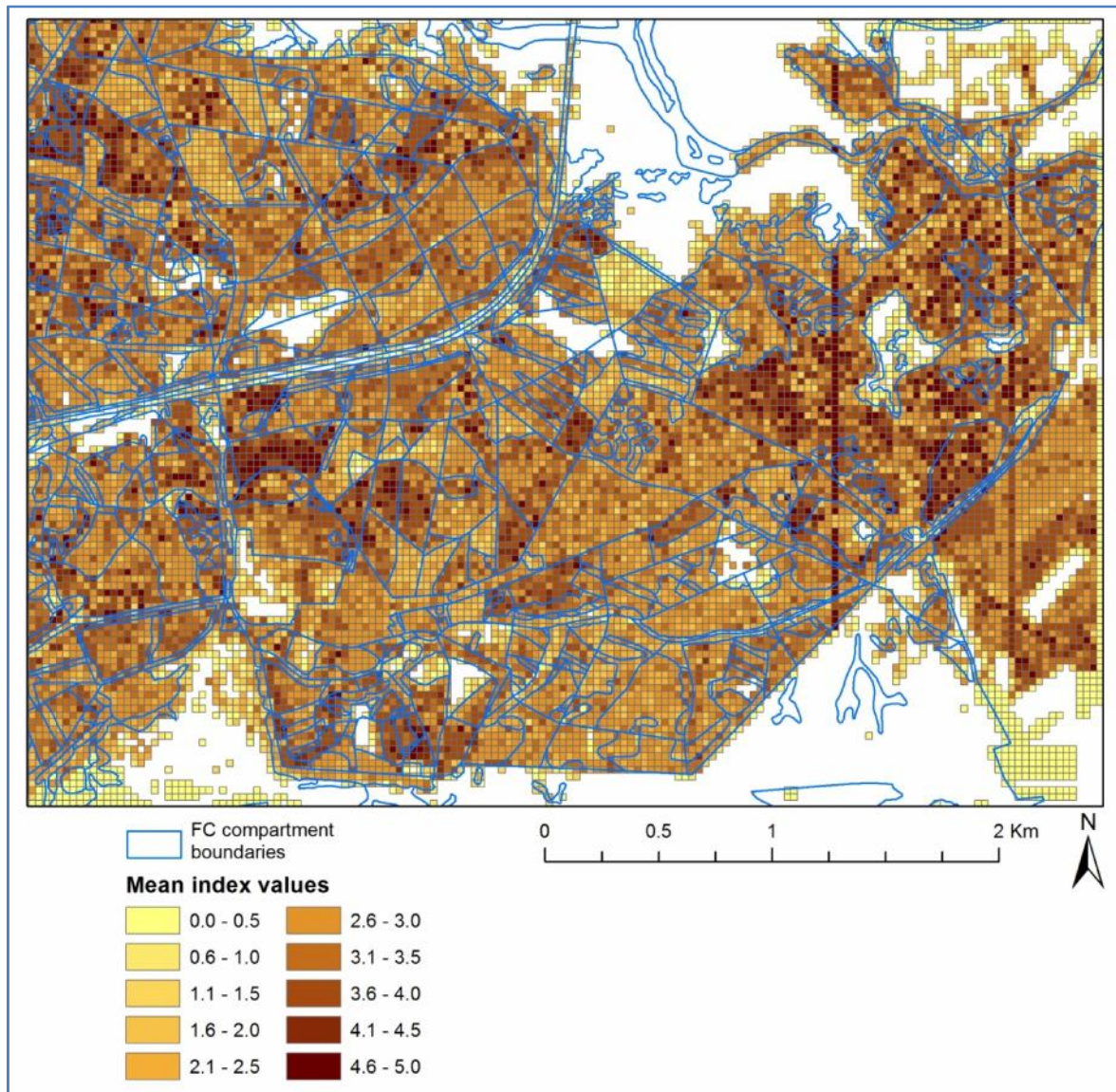


Figure 9.51 – Map of the mean of all three standardised condition indices (HC, CSDI and the score-based index).

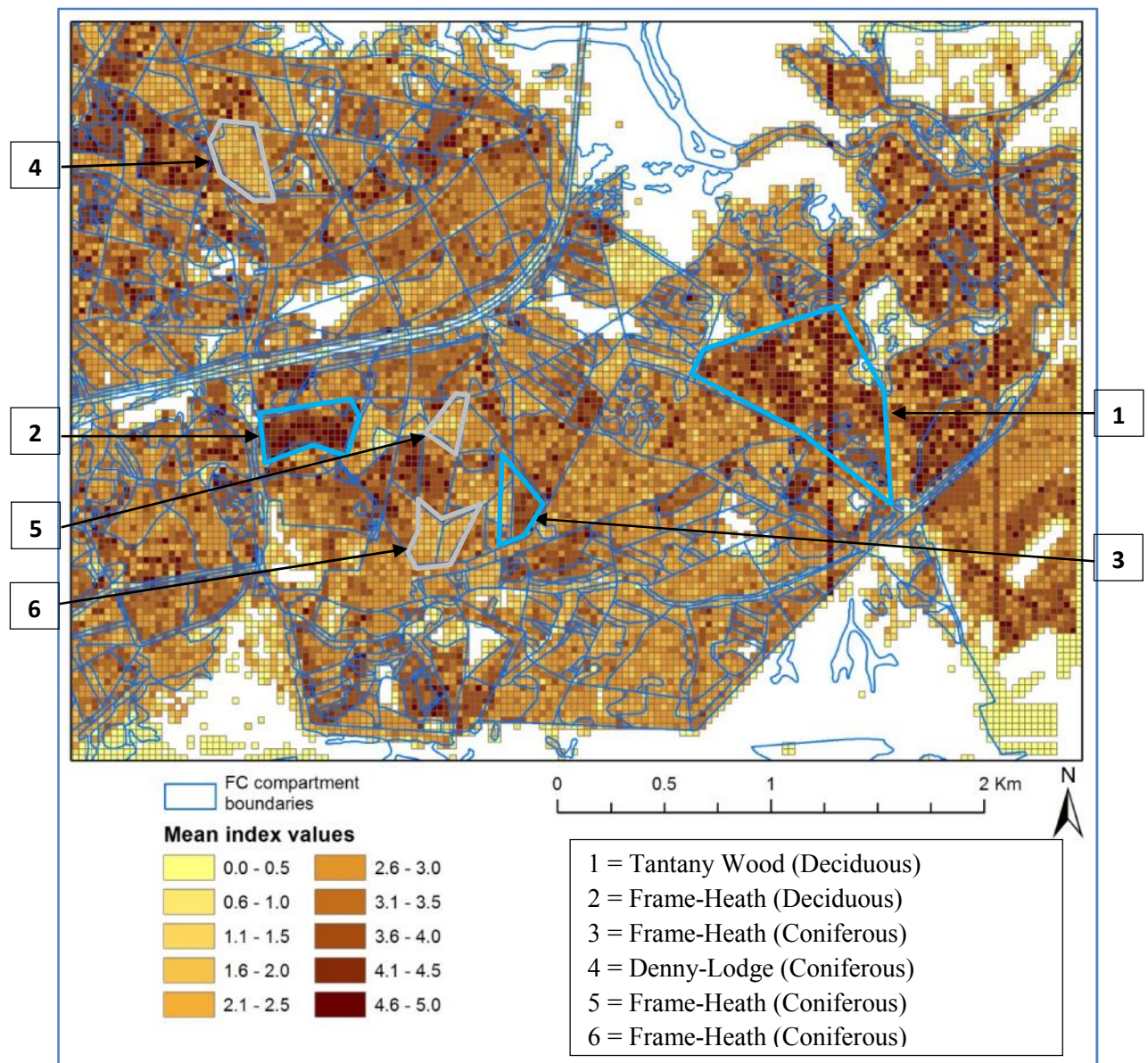


Figure 9.52 – Map of the mean of all three standardised condition indices (HC, CSDI and the score-based index). Annotations illustrate the areas with high mean values.

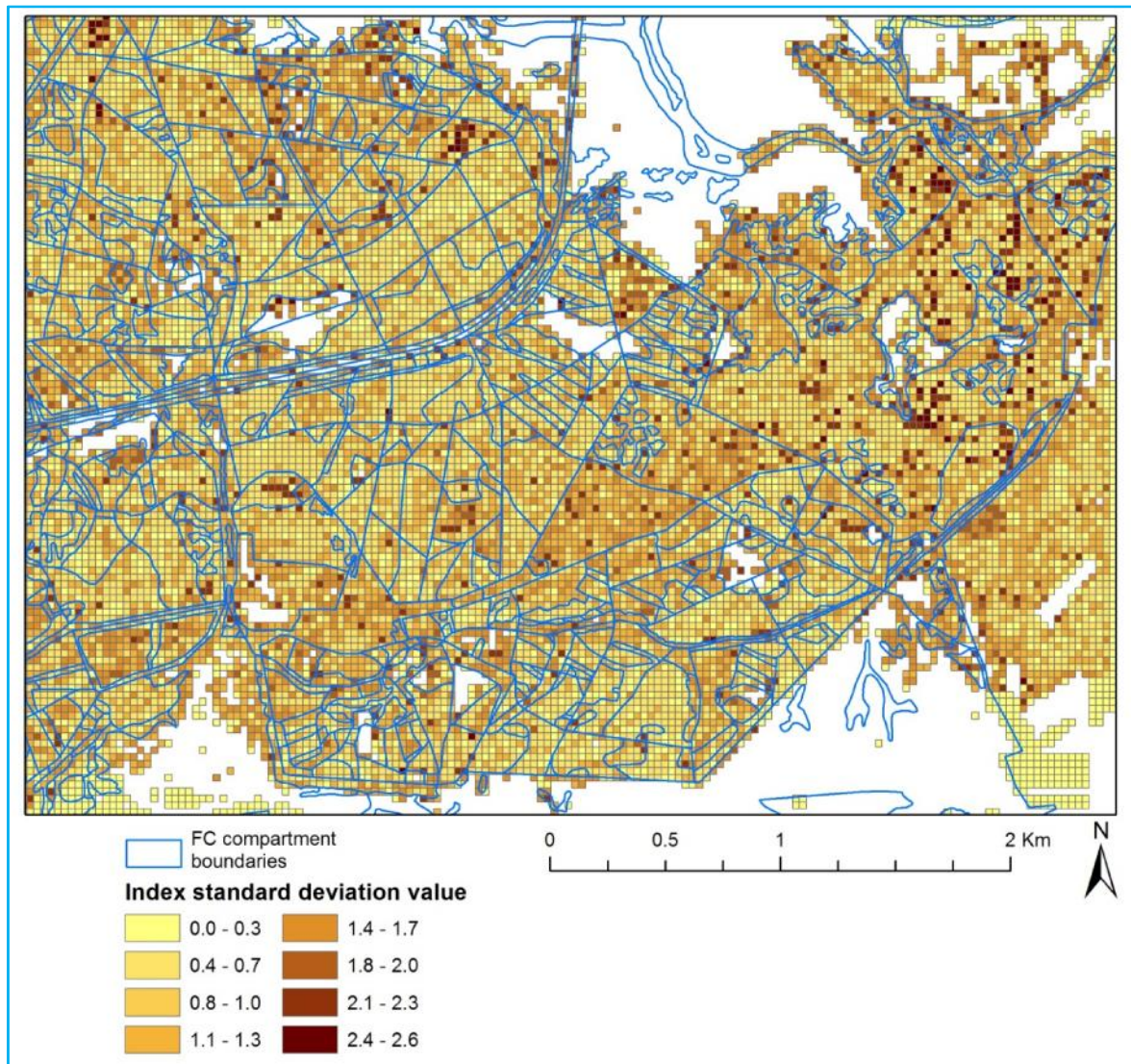


Figure 9.53 – Map of the standard deviation of all three standardised condition indices (HC, CSDI and the score-based index).

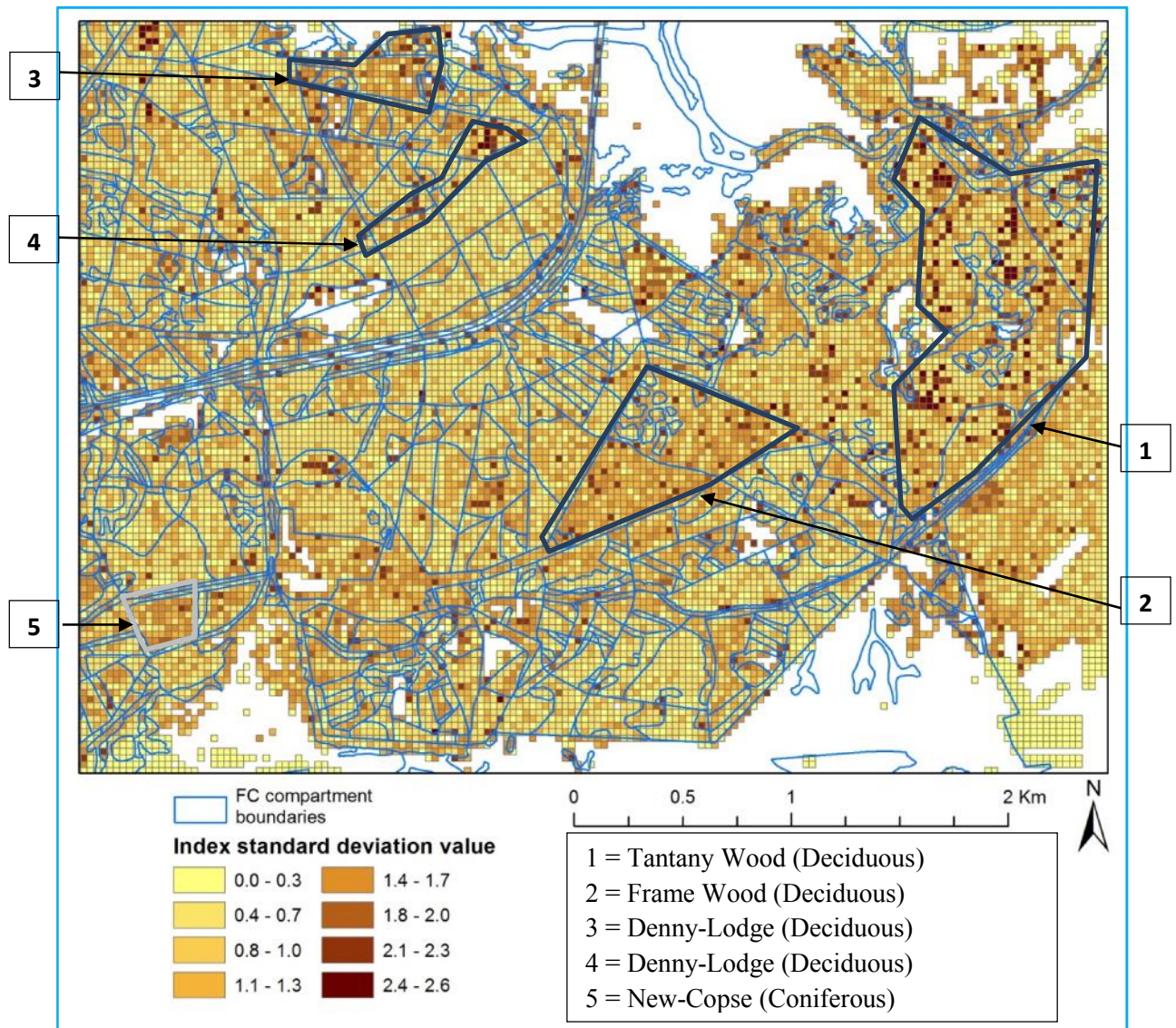


Figure 9.54 – Map of the standard deviation of all three standardised condition indices (HC, CSDI and the score-based index). Annotations illustrate the areas with high standard deviations.

9.4 Summary of findings

All of the 35 successfully predicted field plot-level metrics (identified in section 8.4) were mapped across the study site. The maps presented illustrate the differences in the field plot-level metrics across space and woodland types.

Once the metrics were calculated, the condition index calculations could be applied, for both validation field plots and across the whole study site. Within this chapter, the results of a total of six conventional forest condition assessment indices were presented. Remote sensing derived inputs varied for the indices, employing both metrics extracted directly from the data and derived from statistical modelling approaches. The first of these indices was the VE index, which attempts to describe the horizontal cover within the forest at four vertical levels. A RMSE value of 0.29 (NRMSE 0.29) was returned, and for the majority of cases the remote sensing estimate was similar to that calculated from field data. A number of plot sites' index values were underestimated by the remote sensing calculation; these were in areas of high stem density, where ITC methods had problems in terms of stem counts.

The Clark-Evans aggregation index was calculated using remote sensing derived ITC data. The NN distance was extracted and corrected for each of the tree pairs within the plot extent. A RMSE value of 0.236 was recorded (NRMSE: 0.254). Many of the field calculated indices tended towards the value of one, which was reflected in all but five of the remote sensing calculated indices. Of the remaining five, remote sensing predictions overestimated the field index value. Upon mapping the remote sensing derived index values across the study site it was clear that the majority of the index values occupied the range between 1.29 and 2.15. This very similar range of index values could only broadly distinguish between coniferous and deciduous compartments.

The tree diameter differentiation index was computed using two statistically derived inputs; the first was the estimate of the number of tree stems and the second was the estimate of the sum of the difference in DBH values between stem pairs within the plot. The remote sensing derived index produced a RMSE of 0.189 (NRMSE: 0.232). There was one example where no index value was produced. One plot, number 14, suffered large underestimation from the remote sensing derived index. When the remote sensing derived index values were mapped across the study site the majority of the grid-cell values occupied the range between 0.8-1.0.

Again, this very similar range of index values could only broadly distinguish between coniferous and deciduous compartments.

The complexity index (HC) was calculated using four statistically derived inputs: (i) number of tree species; (ii) the number of tree stems; (iii) the average canopy height; and (iv) the total basal area within the plot. There was a great deal of variation between sites and index predictions from the remote sensing derived inputs. The remote sensing estimates produced a RMSE value of 113.84 (NRMSE: 0.21). The index value for plot 30 was drastically underestimated; this particular plot was a very young, mixed species plot with a high stem density. The map of the HC across the study site for the derived remote sensing index values exhibited spatial patterns which conformed to the different FC compartments.

The CSDI is composed of four indices related to the relative populations of tree species, stem spacing, DBH differentiation and differences in crown size. The remote sensing predictions of index values produced a RMSE of 3.59 (NRMSE: 0.44). Higher field calculated index values were well predicted, generally within ± 1 for the best prediction. Lower field index values were overestimated by the remote sensing prediction by up to 5. The remote sensing map of the derived CSDI value exhibited spatial patterns which conformed to different FC compartments. While there was a great deal of variability in values within semi-ancient woodland, coniferous compartments showed much less variability within their extents.

The four individual index components which make up the CSDI were reviewed. CSDI-1 was based upon the number of tree stems and the relative proportions of the most and least abundant species within a plot. RMSE for this index was high, 0.87 (NRMSE: 0.70) for the remote sensing derived index. CSDI-2 was derived statistically (see Chapter 8.4), the best estimate produced a RMSE of 0.13 (NRMSE: 0.26). CSDI-3 contrasted the three smallest and three largest NN stem distances. The RMSE values for small distances were 1.66-1.90m and 2.15-3.34m for large distances. For the index itself, the subset of best models produced a RMSE of 0.08 (NRMSE: 0.45). The final component, CSDI-4, was constructed by comparing tree crown horizontal extent data and the lowest crown base height. The index calculated from derived remote sensing metrics produced a RMSE of 1.98 (NRMSE: 0.35).

The final index reported was the target and accumulative scoring (or score-based) index, which was constructed through the use of 17 indicator metrics. If the indicator values for the

plot extent exceed the target value, it is scored as one; otherwise it is scored as zero. The sum of these scores is the index value. The index calculated from derived remote sensing metrics produced a RMSE of 2.51 (NRMSE: 0.25). When the remote sensing derived score-based index values were mapped across the study site spatial patterns were observed which corresponded with the FC compartments. The grid-cells which intersected with semi-ancient deciduous woodland exhibited larger index values than those in coniferous compartments. Again grid-cells within coniferous compartments showed much less variability in index value within their extents.

A simple comparison of the mapped index values was carried out by categorising the index-value range, and comparing spatially the products of mean and standard deviation of all index values within each grid-cell. The VE, Clark-Evans aggregation and diameter differentiation indices were removed from this consideration due to uncertainties. Only the HC, CSDI and score-based indices were thus considered. Each of the three indices behaved very similarly within the New Forest study site extent. Each index gave higher values to deciduous woodlands due to the presence of larger trees and trees of more variable size, and lower values to coniferous species dominated compartments. The presence of deadwood seemed to correspond with high values for both tree size and their variability, the presence of understorey, however, was not as closely linked to the aforementioned parameters.

It is obvious that the three condition indices (HC, CSDI and score-based) reflect different elements of the composition and structure of the woodlands in question, especially when applied to deciduous woodland. While there is some overlap between the three, the HC index tends to skew towards high or low values for deciduous or coniferous woodlands respectively, the CSDI responds to differences in structural components irrespective of woodland type, but has issues when canopy gaps are encountered. The score-based index combines elements from the other two in addition to many others, and avoids some of the issues inherent in using ITC metrics directly, which overall provides the most representative condition assessment method.

Chapter 10 – Discussion

The following chapter concerns the examination of the literature search findings and results of this research project. The discussion will be made in relation to the aims and objectives stated in Chapter 1, in particular objectives 3 to 6. The sections in this chapter will discuss the following:

1. a review of hyperspectral derived condition outputs
2. a review of DR LiDAR derived condition outputs
3. a review of FW LiDAR derived condition outputs
4. a review of the best combined hyperspectral and LiDAR derived condition outputs
5. a review of remote sensing derived condition indicators and indices
6. implications and future work

Forest inventories, estimation of structural attributes, biomass and species composition have all been difficult to achieve for large areas through plot based fieldwork measurements alone, because of logistical constraints (Lucas et al., 2008b). High spatial resolution airborne sensors allow imaging of a broad range of features in human related scales. But the advantages of high spatial resolution data, where sample sizes are significantly smaller than the average sizes of the objects of interest, often comes at the expense of an overload in spatial details (Campbell and Wynne, 2011). The following documents the potential of such systems to provide meaningful information.

10.1 A review of hyperspectral derived outputs

The assessment of biodiversity at the local (field-plot) and regional scales often relies on fieldwork-based data collection. Species assessment in relatively large areas has always been a challenging task, mainly because of the difficulty of judging the completeness of the resulting species lists and in quantifying the sampling effort (Palmer, 1995). Inventory of species over large areas is complicated by the fact that every individual within a region cannot be inspected, in addition to species compositional changes over time (Palmer et al., 2006). The following sections discuss the implementation of tree species classification methods and the statistical estimation of forest compositional and structural components from hyperspectral remote sensing techniques to provide answers to the questions set in the project objectives 3, 5 and 6.

10.1.1 Species classification using hyperspectral data

Both the leaf-on and leaf-off hyperspectral datasets required a large amount of processing and experimentation in order to determine the best hyperspectral layers to use in terms of applying dimensionality reduction algorithms and in species classification methods. Within this research project both pixel-based and object-based individual tree crown (ITC) tree species classification approaches were attempted. The pixel-based classification approach was abandoned due to many incorrect and/or missing classes because of dissimilarities in illumination within the individual tree crowns.

It would appear that conventional ‘hard’ per-pixel classification techniques are not sophisticated enough to handle the complexities of vegetation communities in complex forested environments (Levick and Rogers, 2006). Although not part of this thesis, the use of a pixel-based supervised nearest neighbour classification was unable to separate different surface reflectance from canopy and sub-canopy vegetation. Additionally, misclassification of shadowed crown areas was detected. Thus, the object-based classification approach was implemented due to these poor results.

The leaf-on DR LiDAR derived CHM was used to delineate the ITC objects for species classification. These ITC objects were classified using minimum noise fraction (MNF) hyperspectral data which employed a hierarchical classification with user defined membership functions. A more in-depth discussion of the ITC delineation method is presented in section 10.2.1. However it should be noted that the number of ITC objects was smaller than the number of trees found during fieldwork operations for the plot areas and there were a number of over- and under-segmented objects which caused issues with the overall classification.

The classified ITC mapping accuracy (MA) was calculated in comparison to the Forestry Commission (FC) compartment information. The FC compartment information contained up to three tree species classes and their relative proportions as a percentage for the compartment area. The horizontal cover of each of the classified ITC objects was summed by species type and an estimate of map accuracy was calculated. An overall accuracy of 51% was achieved. It should be noted however that fieldwork results could identify more than three species, indicating the basic, and inaccurate, nature of the FC compartment information.

Estimates of MA were calculated for the object-based classification of areas intersecting with field plot extents in two ways. The first concerned a comparison of counts of ITC objects of particular classes and the equivalent for overstorey tree stems. The MA estimated was 49% for the 26 classes. Commission error was an average of 14%, and omission error was 49%. The second method concerned a comparison of the total relative horizontal area of ITC objects of a certain class against the relative area of the field measured horizontal crown area of each species class within the field plot extent. The overall MA calculated for the classification was 62%. Commission error was 27% and omission error was 26%. The latter method was considered the most appropriate.

The levels of classification MA for the relative area assessment approach is comparable in accuracy to the research conducted by Heinzl and Koch (2013) who reported an overall accuracy of 64% for classified ITC objects using hyperspectral data for four tree species classes for a plantation forest in Germany. Lucas et al. (2008b) classified dominant tree species only (eight classes) using classified ITC objects with an accuracy of approximately 80%, in Queensland, Australia. Given the issues inherent in ITC delineation for both coniferous and deciduous forest, and the larger number of species classes present within this study (14 mature and 14 immature tree species), the classification MA is considered comparable.

Owing to the physical shape of tree crowns, different areas of the crown can be more illuminated and other parts can be in shadow, or can be partially covered or shadowed by other trees (Leckie et al., 2005). The attributes of the created image objects through segmentation are based upon the raster cells which make up that object, and the introduction of additional contextual elements such as elevation has definite advantages in aiding classification accuracy. If an optimal segmentation is produced, misclassification of shadowed areas is also no longer a problem. Additionally the classification rules within eCognition allowed the portrayal of gradual boundaries in terms of providing a more realistic interpretation of geographical phenomena (Popescu and Wynne, 2004). The segmentation of canopy areas also has the advantage of focusing the analysis onto features of interest, namely the tree canopy, other objects not fitting the researcher's criteria can be eliminated easily.

Some basic improvements could be implemented into the current workflow to address some the issues present. For example, the selection of the 1.3m height threshold for identifying tree

vegetation could lead to the misclassification of non-tree species, such as gorse (*Ulex*) which can exceed this height threshold. Additional rules and classes within the hierarchy could be implemented to potentially detect and mitigate this issue.

Omission errors were the biggest concern, both in terms of underestimating abundance and the numbers of species, an unfortunate consequence of the automated ITC delineation method, where both individual trees and clusters of trees may be detected. Clusters of trees may mix the input hyperspectral ‘signal’ for multiple trees and cause a misclassification, especially where canopies merge and overlap (Campbell and Wynne, 2011). Similarly over-segmentation within a single tree crown may lead to a misclassification due to the effects of differing illumination of the tree crown (Gougeon and Leckie, 2003; Leckie et al., 2005). An alteration of the automated ITC delineation method is discussed further in section 10.2.1.

There is variability in the spectral signatures between trees of the same species, and because of the effects of tree health and of shadowing within the canopy (Leckie et al., 2005), which further complicate classification of species type. Objects were classified using dimensionally reduced hyperspectral products, namely leaf-on and leaf-off minimum noise fraction (MNF) layers. Layer mean and maximum pixel values within ITC objects were used as part of the membership rules for species classification. A number of studies have extracted spectral values from the individual pixels encapsulated within each ITC object for the purposes of classification, or more specifically, values from the sunlit area for each tree crown rather than all pixels within the object (Bunting and Lucas, 2006; Gougeon and Leckie, 2003; Leckie et al., 2005). Lucas et al. (2008b) utilise a method to identify over-segmented objects based upon the relative position of the maximum pixel value in relation to the object boundary.

The classification within eCognition may also be improved through the use of object attributes related to shape, size or the attributes of neighbouring objects in the identification of over- or under-segmented objects (Trimble Navigation LTD, 2012). If a sufficiently fine spatial resolution LiDAR dataset is available, other DEM products such as the slope and aspect (slope facing) could be used in order to determine area of the tree crown in which to extract the spectra, or alternatively, the selection of specific spectral bands (Leckie et al., 2005). This could also potentially aid in producing a method of removal of over-segmented objects.

For each validation field plot the dominant species type could be identified in all 20 cases. The total number of tree species for each plot could be estimated with an RMSE of 1.8 (NRMSE: 0.31). The relative proportion (percentage) of native tree species detected within each validation plot produced an RMSE of 22.5% (NRMSE: 0.22). The total number of species and proportion of native species provided acceptable levels of accuracy. Conventional methods of estimating species richness within an area, in the form of SH and SI indices could be calculated directly from the ITC objects with an RMSE of 0.56 and 0.28 (NRMSE: 0.39 and 0.22) respectively. Given the diversity indices small range the RMSE values indicate high levels of potential error. Both of the index calculations are based upon the number of ITC objects corresponding to the different species present in each 30x30m area. Given the ITC underestimation and segmentation issues encountered poor estimates were expected.

Remote sensing derived estimates of species richness using object-orientated methods have not seen much use in the surrounding literature (Rocchini et al., 2010a). To date, it has not been possible to find any relevant examples in the literature for the estimation of species richness through the use of metrics equivalent to the SH or SI index which report any measure of accuracy. For example Oldeland et al. (2010) report statistical model results, but no validation or estimates of error were reported.

10.1.2 Extraction of forest metrics relating to condition using hyperspectral data

The combination of all six classified ITC summary metrics (e.g. number of ITC objects, number of native species ITC objects, etc.) and metrics related to the distribution of values within 25 vegetation indices (VI) (e.g. mean, standard deviation, etc.) were used as potential predictors in two regression approaches to model each of the 39 field plot-level metrics.

Two regression approaches were attempted in order to define the best model, ordinary least squares and Akaike information criterion (with a correction for small sample sizes) (AICc). However, RMSE values were very similar between all of the field metric estimates. It was concluded that there is no clear best regression approach for finding a model for estimating the field plot-level metrics.

Many studies have attempted to estimate forest stand attributes by combining airborne spectral and LiDAR datasets, however very few have utilised solely airborne hyperspectral for this purpose. An example of a study which explores the estimation of forest stand

attributes relating to canopy height parameters and estimation of biomass is given in Hyde et al. (2006) to assess the habitat suitability of areas in California, USA, through estimating canopy height and total biomass levels. Estimating soil attributes has also been explored, such as composition and moisture content (Baulies and Pons, 1995). Much of this research focuses on the application of vegetation indices utilising satellite data, a review of which is available in Lutz et al. (2008).

The links between forest structure and composition to VIs are well known, where relationships have been established for green biomass, canopy light use, vegetative stress, water content, etc. A review of this is available in Treitz and Howarth (1999).

Nineteen of a total of 39 field-level metrics, or ~49%, had NRMSE values below a value of 0.4, whereas only three estimates (~7%) had NRMSE values below 0.3. This included both metrics extracted from the ITC summaries and metrics which were statistically derived using both leaf-on and leaf-off hyperspectral data. Unsurprisingly many of the metrics which were related to structural metrics, such as basal area, or volume of deadwood were estimated poorly. Of the 19 models with NRMSE values below 0.4, the number of trees, percentage of native trees, total number of tree species, the SH index and the SI index were best estimated using ITC summary metrics. The remainder of the 19 estimated metrics utilised regression model relationships between one or more ITC summary or VI metrics. Many of the regression models utilised a combination of overstorey composition metrics derived from the ITC summary and VIs with relationships with overstorey architecture, vegetation density and light penetration through the canopy. This is in agreement with findings in the literature, examples of which follow.

The species classification only contributed a relatively small amount of area-based summary metrics; however these ITC-metrics are used in many of the regression models, for example the SI index of species diversity. Some of these metrics, such as ITC count, were not related to species information. Including more metrics relating to the particular species types or proportions present may provide additional useful metrics.

10.1.2.1 Overstorey

In terms of extracting metrics related to the species and size (i.e. DBH) of the trees within the plot, predictive models typically employed indices related to the amount of greenness, or

chlorophyll, in the canopy, which is linked to both species type and canopy architecture. For example, the two broad species groups of coniferous and deciduous have different index value ranges due to different levels of chlorophyll (Richardson and Berlyn, 2002). According to Gspaltl et al. (2013), light use-efficiency increases with increasing tree crown sizes. Trees of native species are more likely to be of the deciduous type, these typically occur in a lower stem density than coniferous woodlands, thus allowing a larger tree size (i.e. DBH).

Models could be produced to estimate the height to the first live branch. The regression models for these metric used the ITC-summary metrics SI and leaf-on calculated VIs. The SI estimate of species richness may provide an indication of the variability of the height of the first live branch as this will typically vary between species types within an area. The VI from leaf-on data has correlations to the forest woodland canopy architecture and the penetration of radiation into the canopy (Chen, 1996). Therefore the VI could indicate the differences encountered between the broad coniferous and deciduous groups, and even within species groups.

Crown area estimates could be made through the use of VIs for the assessment of stress related leaf-pigments and photochemical reflectance. For the former, anthocyanins are typically present in higher concentrations of vegetation (Gitelson et al., 2001), supporting the assumption that crown areas are smaller in areas of higher stem density. The second VI is related to the estimation of light use-efficiency, and therefore the crown size, because of the area available for photosynthesis (Gspaltl et al., 2013).

10.1.2.2 Understorey

Jennings et al. (1999) state that forest canopy attributes, such as light penetration, are related to conditions favouring the survival and growth of plants within the understorey. The narrowband greenness index is related to canopy leaf-area and architecture (Sims and Gamon, 2002), thus it is reasonable to assume overstorey conditions are linked with the understorey. The number of regenerating tree species in the understorey can be estimated through VIs related to the assessment of stress related leaf-pigments, because as mentioned above, anthocyanins are present in higher concentrations of vegetation (Gitelson et al., 2001), and according to Gspaltl et al. (2013) light use-efficiency increases with increasing tree

crown sizes. Thus stem density and light penetration are important factors for the understorey.

10.1.2.3 Deadwood

Research by Pu et al. (2008) using multispectral data suggests that vegetation greenness is linked to mortality levels. Poulos (2009) utilised Landsat ETM+ satellite data combined with a 30x30m DEM to map forest fuels across the Chihuahuan Desert, Mexico. Four classes of fuels were mapped, incorporating alive and dead biomass. A large component of the fuel classes included fallen and standing deadwood volume. In this context deadwood is linked to elevation, cumulative potential relative radiation, Landsat EMT+ bands 1, 2, 3, 4 and brightness, greenness and wetness spectral indices. This study concluded that spectral characteristics were one of the major influences in detecting forest fuels. Alternatively, deadwood volume was generally higher in deciduous areas, which have different index values and ranges in comparison to coniferous areas.

10.2 A review of DR LiDAR derived outputs

The results of this research project reveal the feasibility of using small-footprint discrete-return (DR) LiDAR data to characterise forest attributes required for condition assessment. Many other research studies exist which attempt to extract forest structural metrics and the identification of tree species. What follows is a review of the methods employed here and a comparison to relevant work in the literature. The discussion here relates to project objectives 3, 5 and 6.

10.2.1 Automated individual tree crown extraction

Automated ITC delineation approaches were employed within two different pieces of software, the first was constructed within eCognition using an approach developed by the researcher, and the second used the proprietary software TIFFS. Both of these approaches utilise a similar method to automated ITC delineation, where in summary the ITC objects were grown around ‘seed’ points, defined as the highest raster cell value located within that region of the CHM (Chen, 2006). In comparison the two implementations behaved similarly when validated against field measured overstorey stem number, where the methods produced an RMSE of 16.58 and 15.52 (or 75% and 77%), for TIFFS and eCognition methods respectively. The ITC underestimation of stem counts occurred more frequently when field measured stem densities were higher, and were owing to the lack of differences in height

values within the input raster CHM or too large a search window size to define points of maximum height, both eventualities resulting in clusters of trees being delineated as one object. Trees with large, closely spaced interlocking crowns, or those with heterogeneous crowns, such as old growth deciduous (Leckie et al., 2003a; 2003b), were poorly represented. The segmentation performed best in the well-spaced, predominantly conifer plantation regions within the tree-canopy, a conclusion supported by Tiede et al (2006).

Chen et al. (2006) employed the same ITC delineation technique by locating individual trees in savannah woodland using DR LiDAR, and indicating that the absolute accuracy of the approach was 64.1%. Koukoulas and Blackburn (2005a) outline a different automated ITC detection approach utilising a number of GIS overlay and morphological techniques, such as generating contours, to extract the locations of individual trees, for a New Forest study site in the UK, which intersected with that of the current research project. The authors report a tree-top extraction accuracy of ~80% for trees located in semi-ancient woodland and less than ~50% for trees within plantation woodlands. Lucas et al. (2008b) delineated forest ITC objects using CASI-2 multispectral data. However the ITC delineation utilised follows the same approach as in this study, but with the delineation algorithm using ‘bright’ or high pixel values instead of height data. The ITC delineation process provided accuracies of approximately 70% for individuals and clusters of trees of the same species.

The accuracies of detecting overstorey trees in this project are comparable with those in similar studies. This is an issue explored in Kaartinen et al. (2012), where many methods attempt to define ITC objects from a raster CHM. Image-based methods have a number of challenges relating to how well the vertical profile can be represented and whether relative amounts of over- and understorey can be quantified accurately (Lee et al., 2004).

10.2.2 Extraction of forest metrics relating to condition using DR LiDAR for direct measurements

A number of metrics could be extracted directly from the DR LiDAR dataset which related to field measurements, whilst others were estimated using regression analysis. These are discussed in the following sections.

The directly extracted metrics relate to estimating average tree height, the number of canopy layers and canopy openness from DR LiDAR. The first of these produced an RMSE within

3.46m for all validation field plots. Coniferous sites are generally underestimated and deciduous sites overestimated. These results are similar to those presented in Gaveau and Hill (2003) where estimations of tree canopy height differed by an average of $\pm 2.12\text{m}$ with a standard deviation of 1.24m . It should be noted that other studies have achieved lower error values when estimating canopy height by using statistical methods, such as in Naesset (2002) and Anderson et al. (2005).

The method of measuring tree heights within the field has a number of potential sources of error, this occurs partly because of the difficulty with ground-based measures with sighting the tops of tree crowns that are expansive, or occur where canopy is dense. Coniferous trees typically have crown shapes which taper to a point which often results in underestimation of height by LiDAR measurement, unless a very high LiDAR sampling rate is used or variation in the collection parameters are accounted for (Lefsky et al., 2002). The calculation of Lorey's average tree height for field data is weighted by basal area, thus the larger the basal area the more influence it has upon the average tree height value. Larger amounts of trees in the lower storey of the forest plot may lower the average height resulting in overestimations of deciduous canopy from DR LiDAR.

DR LiDAR estimates of the number of canopy layers produced an RMSE value of 1.18 layers. Unfortunately the researcher could not find any alternative research sources in which to draw a comparison for accuracy assessment. It should be noted however that the number of canopy layers was estimated visually within the field campaigns.

DR LiDAR estimates of canopy openness (%) produced a RMSE of 13.29%. Canopy openness was best estimated using leaf-on data. This accuracy of canopy openness corresponds well with the following two research projects reported in the literature. Morsdorf et al. (2006) calculated canopy cover, again through a regression approach, for a study site located in the Ofenpass Valley, Switzerland, where canopy cover estimates produced a RMSE of 0.18 (or 18%). By contrast Lee and Lucas (2007) estimated canopy closure (the inverse of canopy openness) for a forest in Queensland, Australia, derived through the use of a regression model with a RMSE of 8.6%.

Crown area is often considered a difficult metric to extract, only a limited number of studies have used it in the estimation of forest structural parameters, for example Hyypä et al. (2001), Popescu et al. (2004) and Maltamo et al. (2004). The method outlined in this thesis produced accuracies of within 1.77m and 2.07m for detecting the smallest crown diameters, and within 2.77 and 4.01m for the maximum crown diameters within a plot. The results produced here are comparable to those employed by other researchers, for example Gill et al. (2000) developed models of tree crown radius for several coniferous species in California, USA, and obtained RMSE values from 0.61 to 1.48m. Popescu et al. (2003) produced ITC-based estimates of crown diameter for both coniferous and deciduous trees in Virginia, USA, producing RMSE values from 1.36 and 1.41m for dominant trees.

10.2.3 Extraction of forest metrics relating to condition using DR LiDAR for area-based (30x30m) metrics

As in section 10.1.2, two regression approaches, OLS and AICc, were attempted in order to define the best model. However RMSE values calculated from each of the two methods were very similar for all of the field metric estimates. There was no clear best regression approach for modelling estimates of the field plot-level metrics.

There is a large amount of surrounding research literature concerning the prediction of forest metrics for operations such as inventory, however many only predict a relatively limited number of metrics (Kaartinen et al., 2012; Lim et al., 2003a; Naesset et al., 2004; Richardson and Moskal, 2011; Hudak et al., 2009; Hyypä et al., 2008). What follows is a brief description of the prediction models produced through the course of this research project and a comparison with similar results found in the literature.

A total of 39 field plot-level metrics were estimated through statistical regression using DR LiDAR input metrics. Of these models a number produced high RMSE/NRMSE values; six were above a NRMSE value of 0.5 and thus were considered poor estimates. These models were for the estimation of the standard deviation of tree spacing, the percentage of very big trees (DBH >80cm), SH index (by basal area), the percentage of native saplings, the number of sapling species and the number of bryophyte species.

A total of 30 metric estimation models produced NRMSE values below 0.4; in addition, 16 of these had NRMSE values below 0.3. Two metrics derived from ITC measurements were comparable or better than statistical estimates of mean crown horizontal area and total crown horizontal area. Many of the regression models utilised a combination of area-based point statistics, intensity distribution values and ITC-summary metrics, for both structural and compositional metrics.

ITC summary metrics are used in 17 of the regression models. The estimates of tree number were typically underestimated by ITC counts per 30x30m area. As a result estimates of tree NN spacing were overestimated because of the lower density of ITC objects. The ITC summary value for the number of ITC objects was only used once for estimating the percentage of native trees per 30x30m metric.

The estimates of average and total crown horizontal area and average tree height, however, closely resembled field measurements. As indicated in Chapter 5.2, crown horizontal area was statistically related with tree size (i.e. DBH) and crown dimensional characteristics. A number of other studies have utilised estimates of crown area or diameter in addition to point height distribution metrics in order to predict forest structural attributes, in particular stand-level estimates of DBH and biomass (Hyypä et al., 2001; Maltamo et al., 2004; Person et al., 2002; Popescu et al., 2004).

10.2.3.1 Structural metric comparison

Many of the metrics related to forest structure have been used in research in the application of forest inventory, especially in Nordic countries, for large area surveys. As a consequence of this wide-scale interest there are a large number of researchers who have published results concerning the estimation of stem density, stem diameter (i.e. DBH), basal area, canopy base height or height to the first live branch, and tree crown area metrics.

The first of these, the estimate of stem density, produced in this research a RMSE of 18 (per 30x30m) (NRMSE 0.24). Lee and Lucas (2007) presented estimates of stem counts with a RMSE of 113 stems per hectare, which translates to approximately 12 stems per 30x30m, using an automated ITC segmentation algorithm and linear regression approach with LiDAR metrics for a study site in central Queensland, Australia. Similarly, Naesset (2002) produced estimates of stem number for 200x200m plots for a forest in Våler, Norway, with an error of

between 28 and 35% using a linear regression approach. This approach did not use any ITC metrics.

The research presented in Naesset (2002) also predicted the metrics of DBH and basal area for three types of forest, these being: young, poor condition mature and good condition mature types. Estimates were derived through regression analysis using only LiDAR derived height metrics (e.g. height percentiles). DBH was predicted with a RMSE of 12% accuracy, whereas basal area was estimated with values from 14 to 21%. A similar study by Naesset (2004) produced estimates of DBH with a RMSE of between 13.5 and 20%, and predictions of basal area between 14.8 and 22.5%. Chen et al. (2007a) however produced estimates of basal area utilising both area-based LiDAR point statistics and ITC metrics (crown area) which had a RMSE value of 20.7%. The results of this research project produced similar RMSE values for the 30x30m areas. For average DBH this was 8.4cm (NRMSE 0.29), and for total plot basal area was 0.83m² (NRMSE 0.26).

The height to the first live branch was estimated in this project with a RMSE of 2.7m (NRMSE 0.25) for deciduous, coniferous and mixed woodland types, which was comparable to estimates produced by other researchers. Anderson et al. (2005) produced a linear regression model to estimate crown base height at the 30x30m level which produced RMSE values between 3.9m and 4.1m using LiDAR point height statistics. The objective of that particular study was to map estimated forest canopy fuel parameters for a forest site in Washington State, USA. Naesset and Økland (2002) also produced estimates of the height to the first live branch in the tree crown for 200m² areas in Østmarka Nature Reserve, Norway. Estimates of the height to the first live branch produced RMSE values of between 25 and 37%. Holmgren and Person (2004) utilised a linear discriminant function to estimate crown base height for a coniferous dominated forest area in Remningstorp, Sweden, producing a RMSE value of 2.82m.

To date metrics related to the average spacing of individual tree stems from one another within an area have not seen a great deal of attention, instead other studies have focused on estimating stem density, position, and height values. Here, the DR LiDAR metric used in the regression model was the variance of all leaf-on point heights within the 30x30m extent. The variance is a measure of how spread-out the 3D points are relative to the mean. Therefore it

follows that there is potentially a relationship between the change in the distribution of points within the 30x30m plot and the distribution of forest elements.

The estimation of average crown area was well predicted using ITC-summaries, however slight improvements were made when including one LiDAR intensity metric, the kurtosis of all values. The kurtosis is a measure of the variability of intensity values. The distributions of intensity return statistics, such as the average and standard deviation are related not only related to the reflective properties of vegetation, but also to the larger scale properties of the forest such as canopy openness, spacing and the type of foliage components within individual tree crowns (Moffiet et al., 2005). This supports the supposition that coniferous and deciduous species can be inferred within the model. ITC-summary metrics for total crown area however were superior to those predicted by regression.

A number of field metrics associated with standard deviations were calculated for tree NN space, DBH, height to the first live branch and crown horizontal area. Regression models were calculated for each of these. Unfortunately there are no comparisons which can be made with the literature.

10.2.3.2 Compositional metric comparison

Compositional metrics, such as the proportions of native tree species and the total number of tree species produced relatively low model RMSE values. For the former, metrics relating to the distribution of points (skewness) from both leaf-on and leaf-off data, in addition to the average ITC area, were employed. Deciduous trees within the study site are typically native species where leaf-off conditions cause drastically different point height and intensity value distributions than under leaf-on conditions. They also generally have different tree crown area sizes in comparison to their non-native coniferous counterparts. Also, as Moffiet et al. (2005) states, different tree species may allow more or less returns from within the canopy, which allows species discrimination.

The estimate of the number of tree species utilised the number of returns from the 50th percentile and a measure of the distribution of laser return intensity values. Larger numbers of tree species within an area are more likely to have different vertical profiles – in terms of LiDAR returns – due to the generally different structures within the tree canopy. The distribution of intensity values from the laser pulses interacting with vegetation, in

combination with the presence or absence of foliage and its arrangement within the crown is linked to difference in species, as in Kim et al. (2009a).

The Shannon and Simpson indices of diversity (by stem number) produced relatively low RMSE and NRMSE values, and similarly employ both elevation and intensity value distribution metrics. However, given the small range of the index values, small errors in its prediction could lead to inaccurate results.

10.2.3.3 Deadwood

All of the metrics related to the detection of deadwood utilised intensity metrics from both leaf-on and leaf-off data for ground and non-ground classified points. The results presented in Kim et al. (2009b) are similar to those presented in this project where dead biomass in a forest context can exhibit different LiDAR intensity values when compared with living biomass. Thus with more deadwood within the plot, area-based estimates of intensity metrics will change. Kim et al. (2009b) produced an overall accuracy of prediction of total dead biomass volume of 56.3% using a combination of area-based point height and intensity statistics, for a study site in proximity to the Grand Canyon, USA. Likewise Pesonen et al. (2008) produced estimates for the volume of standing and fallen deadwood for a site in the Koli National Park, Finland, with RMSE values (modified to be comparable to 30x30m area) of 1.6 m³ and 1.3 m³, respectively. The accuracy levels in these examples from the surrounding literature are lower than the 82% (RMSE 1.36m³) for standing deadwood volume and 70% (RMSE 2.49m³) for fallen deadwood volume presented in this study(see Chapter 7.3.2).

There are unfortunately no comparisons which can be made for the decay class metrics however. Both decay class metrics utilised intensity distribution metrics from close to ground-level and/or all non-ground classified points, which may indicate a relationship between canopy structure and deadwood decay.

10.2.3.4 Understorey

To date the detection of understorey characteristics from DR LiDAR only has been limited to estimating either its presence or absence, amount of biomass, or number of suppressed trees or trees with low heights (Estornell et al., 2011; Martinuzzi et al., 2009; Hill and Broughton, 2009; Richardson and Moskal, 2011). Again, each of these metrics were estimated through

some form of statistical regression, apart from in Hill and Broughton (2009), who outlined an approach for detecting understorey presence/absence by applying overstorey tree species specific height thresholds to LiDAR data, for a woodland in Cambridgeshire, UK. The regression approach employed in Martinuzzi et al. (2009) for detecting the presence of understorey utilised LiDAR metrics from low levels in the vertical stratum, i.e. from 0-2.5m, and ground slope information for a forest in Idaho, USA. The approach in Estornell et al. (2011) for estimating shrub biomass utilised LiDAR metrics relating to the mean, maximum and 80-95th percentiles for a forest site in Valencia, Spain.

The LiDAR metrics utilised in the context of this research project for the field-based metrics concerned with vascular species in addition to the prediction of saplings and seedlings, and their species composition, typically employed both LiDAR height and intensity distribution metrics from throughout the vertical profile. A number of these also employed metrics relating to the slope and surface roughness of the ground DEM.

Other researchers have estimated the densities of small standing trees, which will be compared with the number of saplings detected over the course of this project. Richardson et al. (2011) outlines a method of detecting the density of trees between 5-10m in height for a forest in Washington State, USA, by establishing a regression relationship with DR LiDAR metrics which produced RMSE values between 2.60 and 4.58 stems per 0.04hectares (400m²). These plot extents in Richardson et al. (2011) are approximately 44% of the plot extents covered in this project, which produced RMSE values of 24 stems per 30x30m (900m²), or adjusting for the different areas, approximately 10 stems per 0.04ha. Sapling number was low or non-existent across many of the initial field plots (recorded in 2010) which may have caused issues in producing a valid regression model.

10.3 A review of FW LiDAR derived outputs

This section explores the viability of utilising the outputs from small-footprint full-waveform (FW) LiDAR for predicting forest structural and compositional metrics necessary for the assessment of condition. A limited but growing collection of research publications exists for the employment of FW LiDAR data for the assessment of forest characteristics, and there are more which list the potential benefits of such data. Within the following discussion a number of comparisons are made with DR LiDAR in terms of the initial datasets and derived model

estimations in addition to its context within the surrounding research. The project objectives 3, 5 and 6 are be addressed here.

There are a relatively large number of studies in the surrounding literature documenting the effectiveness of different methods for extracting 3D point information from FW LiDAR data, however there are comparatively fewer studies regarding the operational use of FW LiDAR within a forest context; these will be compared where appropriate with this project.

10.3.1 A comparison between DR and FW LIDAR datasets

The results presented in Chapter 7.1.2 illustrate the differences between the DR and FW datasets. The initial number of laser pulses emitted for the FW data acquisition was equivalent to approximately 40% of the pulses emitted for the DR acquisition. Through the detection of multiple returns approximately 30% additional returns were generated for DR data relative to the initial number of pulses, whereas for FW LiDAR an average increase of 143% was recorded in the number of points relative to the initial number of pulses. The Gaussian decomposition method applied to the FW LiDAR data produced similar numbers of total points to that of the DR LiDAR data overall. Similar distributions of points were present within the forest canopy for both types of data, but the FW datasets generally produced a higher number of returns within the mid-canopy.

The different pulse rates between the DR and FW LiDAR datasets influences the 3D spatial sampling and point density beneath the forest canopy. DR LiDAR had a higher density of spatial sampling overall, whereas the FW LiDAR had more points vertically; however there was a larger horizontal spacing at all vertical levels. This will influence the metrics extracted between datasets, however to what degree is unknown.

There was a difference in the sampling resolution for the LiDAR waveforms acquired under 2010 leaf-off and leaf-on conditions, the former using 2ns and the latter using 1ns. The difference in sampling could potentially cause differing sensitivities to features recorded in the returned waveform. It is unknown how the change in sampling may have affected the metrics extracted between leaf-off and leaf-on datasets.

It should be noted that the initial FW LiDAR dataset was larger, by approximately 10 times, in terms of data file size when compared to that of DR, in addition to requiring approximately

double the total pre-processing time required for the DR data. Also, the FW LiDAR processing software, SPDLib, is still under development and may provide additional capabilities in the future.

10.3.2 Extraction of forest metrics relating to condition using FW LiDAR for direct measurements

Similarly to DR LiDAR, a number of metrics could be extracted from the FW LiDAR dataset which could be related directly to field measurements, whilst others were estimated using regression analysis; these are discussed in the following sections.

Metrics related to estimates of dominant tree height per plot, the number of canopy layers and canopy openness could be made directly from the processed FW dataset. All of these metrics were calculated based on the distribution of height measurements, and as such are comparable with those made with DR LiDAR.

The estimates of dominant tree height produced RMSE values of 3.7m, which was similar to DR estimates (RMSE 3.46m) overall (see section 10.2.2). This is again comparable to measurements presented in other studies. For example, in Gaveau and Hill (2003) estimations of tree canopy height differed by an average of 2.12m with a standard deviation of 1.24m. Again, coniferous sites are generally underestimated and deciduous sites overestimated. The difference between the two datasets may be attributable to the difference in sample densities from DR and FW, as the latter had a much smaller overall pulse density, canopy height differences may have been missed between acquisitions.

Prediction of the number of canopy layers using FW LiDAR produced RMSE values of 1.07, while similar to DR estimates (RMSE 1.18), they were slightly lower and more closely resembled the visual field assessment. Likewise FW prediction of canopy openness produced RMSE values of 8.28% for leaf-on data, which again is a slight improvement in comparison to the DR leaf-on estimate (RMSE 13.29%). As mentioned previously, Lee and Lucas (2007) estimated canopy closure (the inverse of canopy openness) through the use of a regression model with an RMSE of 8.6%.

When comparing DR and FW outputs, it is possible these decreased error values is because of the increased number of points resolved within the mid-canopy, through FW waveform

decomposition, which is an observation supported by Reitberger et al. (2008). Given the overall similarity of the DR and FW datasets, however, the differences in extracted statistics were very small.

10.3.3 Extraction of forest metrics relating to condition using FW LiDAR for direct measurements area-based (30x30m) metrics

The estimation of field-level area-based metrics from FW LiDAR data inputs was performed using statistical regression approaches. OLS and AICc methods were used in order to predict the metrics; there was no definitive best approach, with both often performing similarly. Out of a total of 39 metrics, five produced high RMSE and NRMSE values (i.e. $\text{NRMSE} \geq 0.5$). These five metrics were: the mean DBH, the percentage of very big trees (i.e. $\text{DBH} > 80\text{cm}$), the SH index (by basal area) the percentage of native saplings, the seedlings SH index for native species and the number of bryophyte species. DR LiDAR estimates of these values, apart from model estimates of mean DBH and the seedlings SH index for native species, could also not produce satisfactory predictions.

A total of 26 metric predictions had values of NRMSE below 0.4; of those a total of 18 had NRMSE values below 0.3. As with the DR LiDAR, FW LiDAR ITC-summary estimates of mean crown horizontal area and total crown horizontal area were comparable to the statistical estimates, the latter of which exhibited a slightly lower error value for the ITC derived metric. Many of the regression models produced used combinations of ITC-summary, height, amplitude and width metric. It should be noted that the amplitude and width metrics were utilised to a much higher degree than that to which intensity was used in the DR derived models.

A total of 17 FW LiDAR derived models employed at least one of the ITC-summary metrics in its construction. This is in common with DR LiDAR derived models, where estimates of crown diameter in particular were important in many of the predictive models, this is supported by existing research (Hyypä et al., 2001; Maltamo et al., 2004; Person et al., 2002; Popescu et al., 2004).

The decomposition of the individual waveforms allows the extraction of additional metrics per point compared with DR LiDAR. In particular, the metrics related to peak-width can be extracted for each derived 3D point, i.e. the width of each peak fitted to each of the surface

interactions (i.e. returns) in the digitised return waveform. According to Lin and Mills (2010) the width metric is related to the slope and surface roughness of the material the laser pulse interacts with and exhibits relatively stable behaviour when amplitude, range distance or scan angle vary substantially. Wagner et al. (2006) however state those echo-width estimates at high amplitudes are relatively stable, but show significant scattering effects at lower amplitudes. It should be noted that in Figure 4.15 anomalies in echo-width values were detected when the FW LiDAR flightlines were merged, with higher values present on these overlapping areas which were furthest from nadir, which contradicts the above statement. This is why metrics recorded furthest from nadir were removed, reducing this problem.

FW LiDAR metrics related to amplitude only characterise the peak power of the returned echo (Wagner et al., 2008). This metric is comparable to DR intensity. According to Alexander et al. (2010) amplitudes of echoes from targets available from airborne FW data have been found to be useful in the identification of different forest cover types. Amplitude is however dependent on various factors such as range and incidence angle.

10.3.3.1 Structural metrics

FW LiDAR estimates of forest structural metrics produced similar results in terms of RMSE and NRMSE values to the DR LiDAR derived estimates for the commonly extracted forest metrics of: stem density, stem diameter (i.e. DBH), basal area, canopy base height or height to the first live branch, and tree crown area metrics. These models typically employ both leaf-on and leaf-off datasets.

The first of these is stem density, which in the context of this project had an RMSE of 16 (per 30x30m) (NRMSE 0.22) which slightly lower than the DR predicted equivalent (RMSE 18, NRMSE 0.24). The results were still in line with the accuracies provided by other sources, for example Lee and Lucas (2007) with an RMSE of approximately 12 stems per 30x30m. Likewise estimates of stem spacing were very similar between FW and DR data, where RMSE was 1.3m and 1.1m, respectively.

FW LiDAR estimates of plot mean DBH value were above NRMSE 0.5 and were inferior to the DR LiDAR equivalent. FW LiDAR model predictions for total plot basal area produced RMSE values of 0.9m² (NRMSE 0.28), which were very similar to DR LiDAR estimates (RMSE 0.83m, NRMSE 0.26). The values produced for this project have a slightly higher

error than those presented in Naesset (2002, 2004) where RMSE values range from 14 to 22.5%.

FW LiDAR regression model estimates for height to the first live branch (or canopy base height) produced an RMSE value of 2.6m (NRMSE 0.24), again which was similar to those produced for DR LiDAR (RMSE 2.7, NRMSE 0.25). As mentioned previously this compares favourably to the estimates of height to the first live branch presented in other studies, for example that which is presented in Holmgren and Person (2004) with an RMSE value of 2.82m.

ITC crown area metrics were incorporated in the FW regression model for the prediction of average crown horizontal area, which also included metrics relating to the above ground amplitude value distributions. FW model estimates produced a RMSE value of 24.4m^2 , which is very similar to the estimate developed for DR LiDAR where the RMSE value was 23.5m^2 . The total crown area was best predicted by the ITC metric; in this regard the FW and DR estimates were identical.

A number of field metrics associated with standard deviations were calculated for tree NN space, DBH, height to the first live branch and crown horizontal area. Regression models were calculated for each of these. Unfortunately there are no comparisons which can be made with the literature. The first of these, the standard deviation of NN space for DR LiDAR produced a RMSE of 0.57m, whereas FW estimates produced a lower RMSE value of 0.39m. The standard deviation of DBH produced an RMSE of 10.57cm from DR LiDAR. The FW estimate produced a smaller value of RMSE, at 9.1cm. The standard deviation of height to the first live branch gave an RMSE of 1.66m from DR LiDAR while that produced using FW LiDAR was higher, at 2.22m. The standard deviation of crown horizontal area gave an RMSE of 20.62m from DR LiDAR. The FW LiDAR estimate gave a lower RMSE value of 17.59m.

10.3.3.2 Compositional metrics

The initial compositional metrics relating to the proportion of native trees and the number of tree species within the plot produced relatively low RMSE values using FW LiDAR regression models. In the comparison of FW and DR metrics, RMSE values were often

similar; the former better estimated the number of native trees and the total number of tree species, but the percentage of native trees was best modelled with DR LiDAR.

Even the Shannon and Simpson indices of diversity (by stem number) produced relatively low RMSE and NRMSE values, and similarly employ both height, amplitude and width value distribution metrics. However, given the small range of the index values, thus errors in its prediction could cause drastically different values to those derived from field data.

Each of the compositional metrics was modelled using combinations of height, amplitude and echo-width attributes, apart from the model to estimate the percentage of native trees. Metrics relating to heights are indicative of vertical structure (Reitberger et al., 2008). Amplitude can be considered an indication of the material the laser pulse interacts with, although variability is introduced with range and scanning angle (Wagner et al., 2006). The echo-width metric, of the returned ‘echoes’, has been found to be an indication of surface roughness and slope angle of the material the laser pulse interacts with (Lin and Mills, 2010). It can be assumed that variation in these metrics is indicative of changes in vegetation composition, for example the associated structural differences between species.

This assumption is supported by the research presented in Reitberger et al. (2008) where a method is outlined for classifying ITC objects using a combination of leaf-on and leaf-off acquired FW LiDAR data. The authors reported that amplitude does not distinguish well between coniferous and deciduous trees under leaf-on conditions, but it is better, however, in distinguishing between bark and coniferous-needles in leaf-off situations. They also state that combining geometric, amplitude and width metrics improved accuracies of species classification, and noted in particular the usefulness of the echo-width metrics. Likewise, Heinzel and Koch (2011) were able to classify six tree species types by combining amplitude, echo-width and point density metrics with accuracies of 78% for coniferous and 91% for deciduous species for a site in south-west Germany.

10.3.3.3 Deadwood

At the time of writing this thesis there were very few examples in the surrounding literature of estimating forest deadwood volume using FW LiDAR data. One such example is the work presented by Mücke et al. (2012) which focuses on the detection and mapping of large fallen deadwood items. The estimation of standing deadwood using FW LiDAR data produced

RMSE values of 1.36m^3 (NRMSE 0.17). This estimate was again similar to that derived through the DR LiDAR, where the RMSE value was 1.50m^3 . Metrics related to the variation in ITC object spacing, the leaf-on echo-width at lower height levels within the plot and the distribution of leaf-on non-ground amplitude were used as model predictors. Estimates of downed deadwood volume using FW LiDAR data produced a RMSE value of 2.5m^3 . In comparison, DR LiDAR estimates produced a RMSE of 2.7m^3 . The metrics employed in the model included an indication of the distribution of the non-ground intensity and the variation in the leaf-on heights within the plot. With the exclusion of the echo-width metric, this is very similar to the input metrics with DR LiDAR, where intensity (i.e. amplitude) is linked to total biomass volume (Kim et al., 2009b; Martinuzzi et al., 2009).

Standing and downed deadwood decay class estimated from FW LiDAR produced RMSE values of 0.19 and 0.20, respectively, for estimates on a scale of 0-1. In comparison, DR LiDAR estimates were similar, producing RMSE values of 0.24 and 0.17, respectively. Both decay class metrics utilised amplitude distribution metrics from non-ground classified leaf-off points. Standing deadwood decay class utilised height percentiles whereas downed deadwood decay utilised echo-width percentile metrics. As with the DR models, this may indicate a relationship between upper canopy structure and deadwood decay class. This is supported by the correlations between the plot-level deadwood metrics and crown horizontal area and height to the first live branch metrics found in the bivariate correlation analysis in Chapter 5.2.1.

10.3.3.4 Understorey

As with studies into the detection of forest understorey structures and individual plants for DR LiDAR, there is only a small amount of literature concerning this for FW LiDAR. Of the limited number of studies found, they focus upon the detection of understorey trees through 3D point cloud segmentation (Reitberger et al., 2009).

There are a number of directly comparable metrics to studies which utilise DR LiDAR, for example the estimation of the densities of small standing trees, with those trees classified as regenerating in this current project (i.e. saplings). Richardson et al. (2011) produced RMSE values between 2.60 and 4.58 stems per 0.04 hectare (400m^2) areas. These plots are approximately 44% of the plot areas covered in this project, which produced RMSE values of 26 stems per $30\times 30\text{m}$ (900m^2), or adjusting for the different areas, approximately 11 stems

per 0.04ha. In comparison to DR LiDAR within this project, estimates of saplings produced RMSE values of 24, which was very similar to the FW estimate. Sapling numbers were low or non-existent across many of the initial 2010 field plots which may have caused issues in producing a valid regression model.

The composition and structure of the forest understorey metrics estimated through FW LiDAR produced similar overall results to those by DR LiDAR. Regression models derived from both datasets could not estimate the percentage of native saplings or the number of bryophyte species.

The number of saplings and number of native saplings have very similar FW LiDAR metric model inputs involving low to medium height metrics for leaf-on and leaf-off metrics and echo-width metrics for low heights from leaf-on data. Saplings recorded during fieldwork generally were of small stature, and thus it makes sense that the required metrics would come from this part of the height stratum.

The number of sapling species was only modelled successfully when derived by FW LiDAR, giving an RMSE of 1.02species per 30x30m. It should be noted however that the number of saplings across the entire study site was very low which may have caused issues in the predictive regression models. In contrast, only DR LiDAR could predict the Shannon index for native seedlings species, the RMSE value was high (0.72). The small and defined range of the index value is also a concern.

The estimation of the number of seedlings per plot for DR LiDAR produced slightly lower RMSE values than for FW, 295 and 332 respectively. For FW LiDAR, the regression model utilised one input relating to the variation in leaf-on amplitude for all points, however the DR model used four inputs relating to ground slope, and measures of intensity variation in ground and non-ground points.

The number of native seedlings per plot for FW LIDAR model estimates produced a RMSE of 245, and 239 for DR LiDAR. FW metrics again included measures of the leaf-off amplitude variability in addition to both leaf-on and leaf-off measures of ground echo-width. DR LiDAR uses solely ITC-metrics of the variability in ITC space and average canopy volume. Variation in ITC values is typically related to coniferous or deciduous species, where

the latter generally has more native species present and generally larger tree crown horizontal areas. Variation in amplitude and echo-width however has links to different materials and surface slope and roughness attributes recorded within the laser pulse footprint, which in this case would appear to indicate the estimation of seedling properties from direct measurements from ground-level.

The number of seedling species encountered within each plot was estimated by DR LiDAR with a RMSE value of 3.2 stems, whereas for FW LIDAR RMSE was 3.1. DR metric inputs to the predictive model utilised leaf-on values of intensity and height for the 60th percentile, and values of leaf-off canopy relief. This potentially highlights links to the cover and general species types in the mid-canopy. FW metric inputs to the predictive model on the other hand, utilised both amplitude and echo-width values from the ground classified points for leaf-off data only, which appears more sensible as seedlings will occur at ground-level only.

The estimation of the vascular plant species number for DR LiDAR regression model produced a RMSE value of 3.4 species, whereas FW LIDAR model produced lower RMSE value of 1.7. The FW predicative model only required one input, the variation in heights of non-ground leaf-on points. The DR model however required four inputs utilising metrics relating to topographic roughness, low-canopy strata intensity, the distribution of height points and average canopy height. The derived FW model was not as complex and provided a similar accuracy level to the DR equivalent, which indicates a better model.

10.3.4 LiDAR sampling disparity

In the context of this research project FW models could replace DR models with very minor changes in prediction accuracy. As mentioned previously, the DR and FW LiDAR were collected with substantially different initial laser pulse repetition rates, resulting in fewer FW pulses being emitted. Even with the fitting of a greater number of points through Gaussian fitting for the FW LiDAR data, providing a great deal of information along the vector of the laser pulse, the distribution of points and total sampled forest elements will be different to DR LiDAR data. According to the studies of Treitz et al. (2012) and Strunk et al. (2012) the precision of statistical estimates of vegetation structure estimations showed little change when the point density of DR LiDAR data was varied from 3 pulses m⁻² to 0.5 pulses m⁻². Their structural metrics were basal area, timber volume, biomass, stem density and canopy height, thus some of the other metrics mentioned here may hypothetically behave differently.

It should also be noted that these tests were performed using height metrics only, intensity, FW amplitude or echo-width metrics were not addressed. It is therefore unknown whether these other attributes will change with point density.

10.3.5 Concluding thoughts

Overall, it would seem that the models produced using FW LiDAR data were comparable to DR LiDAR, and overall RMSE values are similar. The initial dataset for FW LiDAR equated to 40% of the DR LiDAR. FW LiDAR model predictions presented here compare favourably with those published in the surrounding literature, for both FW and DR LiDAR. Upon exploration of the model inputs, it is clear that the FW derived models make extensive use of echo-width related metrics, an observation also made in Reitberger et al. (2008). It is also evident that the regression models derived through using FW LiDAR often utilised fewer and/or more appropriate input metrics than the DR derived models, especially so for the prediction of understorey composition metrics. In addition, as with DR LiDAR, ITC estimates of crown area metrics were utilised a great deal in model construction.

10.4 Exploration of the fusion of the airborne remote sensing datasets for assessing condition

This discussion section will focus upon reviewing the combination of remote sensing datasets for the goal of estimating fieldplot-level metrics critical to the assessment of woodland condition. Within the course of this project predictive regression models were derived from combinations of hyperspectral and DR LiDAR data and then for hyperspectral and FW LiDAR data in order to test if better metric estimations could be made by the addition of supplementary datasets.

As in the above, two regression approaches were attempted in order to define the best model, these were OLS and AICc. However, RMSE values calculated from each of the two methods were very similar between all of the field metric estimates. There is no clear best regression approach for finding a model for estimating the field plot-level metric.

What follows is a review of the best predictive models (i.e. the lowest RMSE) developed for all of the remote sensing datasets, and relevant explanations of model inputs and comparisons to existing research. The project objectives, in particular 5 and 6, will be addressed here.

The best estimates of the field plot-level metrics from airborne remote sensing datasets came from a mix of hyperspectral, DR and FW LiDAR. In addition, a number of models combined metrics from both hyperspectral and LiDAR. What follows is a review of the predictive models tested which produced the smallest RMSE value overall for that field metric. It should be noted that RMSE values between the two different input dataset groupings were often very similar. Table 10.1 describes each model and explains the metrics utilised in its construction in relation to the surrounding literature. It should be noted that two metrics related to index construction, the sum of diameter differences and the index of vertical separation were not reported in this table.

The individual metric prediction accuracies have been compared to those in the surrounding literature in the previous sections, other than those metrics derived from combining hyperspectral and LiDAR metrics in one model. A full review of all the regression model results produced by the composite datasets is not necessary as many of the model estimations did not show much improvement, i.e. by reducing RMSE. In relation to the best predictive models for all potential datasets however, a total of nine of the composite dataset models, (i) four for hyperspectral combined with DR metrics, and (ii) five for hyperspectral combined with FW metrics, provided better estimates (i.e. lower RMSE values) than for each of the other datasets in isolation. For hyperspectral combined with DR LiDAR these were the SH index (by stem count), the SI index (by basal area), the total crown horizontal area and the number of native saplings. For hyperspectral combined with FW LiDAR these were the number of native trees, number of sapling species, number of seedlings per plot, number of seedling species and the seedling SH index for native species.

There are few examples in the surrounding literature of combining remote sensing datasets relevant to those presented within this thesis. In terms of species classifications, a number of authors have mixed spectral and LiDAR raster layers for land cover classifications. Hill and Thompson (2005) were able to develop a method to combine dimensionally reduced hyperspectral and LiDAR data using unsupervised Isodata classification. Ten classes for canopy and shrub characteristics were produced. Mutluet al. (2008) also utilised fused dimensionally reduced satellite multispectral and airborne LiDAR data to classify land cover in relation to fire fuel risk.

Lucas et al. (2008b) on the other hand utilised a method to predict total above ground biomass. The method combines ITC algorithms for the mapping of overstorey and suppressed tree stems, which utilise inputs from multispectral and LiDAR data, species classification and the application of species specific allometric equations per stem for estimating biomass.

Each of these fusion applications identified in the surrounding literature focuses ultimately on the identification of a single attribute such as the identification of species types. Many of the studies encountered in the surrounding literature, for estimating forest metrics from a single remote sensing dataset for example, are also somewhat limited in the number of attributes they attempt to estimate. This research project in particular has estimated many more compositional and structural metrics within a reasonable level of accuracy (i.e. NRMSE <0.4) across a wide area.

Over the course of this thesis project a fusion of the two datasets types was applied multiple times, for example the species classification of ITC objects delineated by an algorithm acting upon LiDAR data, in addition to the combination of metrics for regression modelling. In order to produce the best overall method of estimating forest structural and composition metrics it was necessary to derive models utilising inputs from each of the three input datasets.

Only two relatively minor metrics could not be estimated, these were the percentage of native saplings and the number of bryophyte species. The percentage of native trees, saplings and seedlings metrics were introduced only as an alternative value should the estimate of stem count not function correctly. The number of bryophyte species could have been useful as it is a required metric for some alternative condition assessment methods (Van Den Meersschaut and Vandekerckhove, 1998). However it was not utilised in any assessment indices within this project.

For the 37 metrics which could be modelled, some metrics could only be extracted from either hyperspectral, LiDAR or a combination. Some datasets alone could not produce estimates at all. The wealth of information the cumulative datasets provided does account for the shortcomings of the individual datasets. However, two metrics produced poor estimates, these were the percentage of very big trees (DBH >80cm) and the Shannon index (by basal area).

The statistical approaches of OLS and AICc regression often produced very similar RMSE values when applied to the same dataset. This often resulted in an approximate 50/50% split of predicted models which used OLS and AICc for all five datasets tested. There is unfortunately no clear best regression type which can be discerned from either of these methods.

In particular, FW LiDAR provided the most input to predicting 20 (15 LiDAR and 5 composite) (57%) of the field-level metrics, in comparison to 13 for DR (9 LiDAR and 4 composite) (37%), and 2 for hyperspectral models (6%). In many cases, the RMSE value for models derived from the five datasets were very similar, in particular this was evident between DR and FW LiDAR.

FW LiDAR provides a benefit in predicting a significant proportion of forest structural and compositional metrics, and has proven to be as good as DR estimates, if not better given the initial disparity between sample densities. This is owed in part to the additional metrics, in particular echo-width, which FW LiDAR provides. It is also interesting to note that both DR and FW LiDAR datasets often provide components of models which estimate forest compositional metrics, such as the number of tree species, which is typically the realm of spectral data alone. This is supported by Hill and Thompson (2005) and Multu et al. (2008) where combinations of hyperspectral and LiDAR data were used to provide forest cover classifications.

A small number of relatively poor estimates were observed when comparing predicted results against those measured in the field. This was typically for very dense mixed stands, such as for plot number 30. None of the initial training field plot sites were similar to this validation plot due to the very high density of young or small trees. This highlights the need for more and appropriate fieldwork for training the production of statistical predictive models. As in Strunk et al. (2012) the number of training samples has a dramatic effect on the validity of inferences which can be made from remote sensing predictive models, as expected from sampling theory. For example, the sapling population was generally low across all field plots; this low overall population may have caused issues in estimating its number.

As mentioned before, the SH and SI indices of diversity (by stem number) produced relatively low RMSE and NRMSE values. However, given the small range of the index

values, caution should be applied when using the results, as even small errors in its prediction could result in different results when compared with field data. Likewise metrics such as the percentage of non-vegetated or bare earth cover were assessed visually and may be subject to surveyor bias and inaccuracy. The number of seedlings was enumerated using a 10x10m sub-plot which was assumed to be representative; this may add additional error to the estimations.

As noted previously, metrics such as saplings were low across the entire study site. It is likely the model produced would perform poorly in areas where sapling numbers were high. Issues such as this in the underlying data could lead the researcher to question how realistic some model estimates are.

Finally, each of the statistical regression approaches utilised in the course of this thesis involved hundreds of remote sensing metrics (>300), many of which exhibited signs of multicollinearity. Although efforts were implemented to mitigate this problem, there were many metrics portraying the same or similar information for both hyperspectral and LiDAR datasets. There are potentially many more possible combinations of remote sensing metrics which could be used to estimate field metrics. In terms of the issue of causality and correlation, some of the models produced may not be directly related to the remote sensing inputs, for example deadwood decay class was modelled using hyperspectral data, which may be related more to VI values linked to being located in deciduous woodland, which universally had high deadwood volume and decay class, rather than a link between VI value and understorey attributes.

There were a number of 30x30m grid cells within the shapefile map which contained large canopy gaps, or very low stem densities. These typically existed on the perimeter of the study site, and in a limited number of locations within the forest which were cleared prior to the data acquisition in 2010. Unfortunately a small number of examples of large metric overestimation were caused by this as all regression training data were acquired from within the forest itself. These grid cell values could be ignored, and in the future, a correction could be applied to remove these cells.

Table 10.1 – The predictive regression models with the smallest RMSE

Field metric	Dataset	Remote sensing metrics description
<i>Number of trees</i>	FW LiDAR	<p>This model used the metric inputs of ITC derived average crown area, a percentile from low in the height stratum for leaf-on data, the distribution of amplitude values and the distribution of height values, both from leaf-off data.</p> <p>Other studies have reported the benefits of using estimates of crown area and height distribution metrics in the prediction of other forest structural characteristics (Hyypä et al., 2001; Person et al., 2002; Popescu et al., 2004; Maltamo et al., 2004). There is also a strong correlation found between average crown area and the number of trees found in the bivariate correlation (section 5.2.1). Smaller crown sizes are generally indicative of higher populations, especially so in plantation woodland. Reitberger et al. (2008) indicates that amplitude metrics can distinguish between tree bark and coniferous-needles for use in species classifications, thus the distribution of amplitude values per plot could be an additional distinction between broad species type. Moffiet et al. (2005) and Kim et al. (2009a) indicate intensity/amplitude summary metrics are also related to attributes such as canopy openness, which is dependent on stem density and species types.</p>
<i>Number of native trees</i>	Hyper. and FW LiDAR	<p>The predictive model utilises two inputs, the first is the leaf-on simple ratio vegetation index (SRI) and the second is a measure of leaf-off echo-width (FW) from within intermediate canopy height.</p> <p>The SRI is a measure of broadband greenness, which has been linked to canopy leaf-area and architecture (Sims and Gamon, 2002). Native tree species are typically deciduous, whereas non-native species are coniferous, which typically have a very different structural form. We can assume the SRI is distinguishing between the two broad species types. The FW metric regarding echo-width is related to surface roughness and slope. Echo-width is useful in discriminating between vegetation cover types (Reitberger et al., 2008; Heinzel and Koch, 2011)</p>
<i>No. of tree species</i>	DR LiDAR	<p>This model utilised two DR LiDAR measures of return intensity, the first was the leaf-off 50th percentile, and the second was the leaf-on variance of non-ground intensity.</p> <p>As in Brandtberg et al. (2003), this model implies that different tree species can be separated by LiDAR intensity metrics. The intensity distribution within the mid canopy may be indicative of the presence of different numbers of tree species within the plot extent, as different tree species can exhibit different canopy architecture and intersect with this level, as in Kim et al. (2009a). The variance of intensity values could also be indicative of differing species numbers where the presence of many disparate species, and thus intensity values would result in a differing distribution of intensity values.</p> <p>It is interesting to note that this DR model achieved higher accuracies than the hyperspectral classification direct measure.</p>

Table 10.1 (continued)

Field metric	Dataset	Remote sensing metrics description
<i>Average tree nearest neighbour(NN) spacing</i>	FW LiDAR	<p>The model employed two input metrics; the first was the spacing of the ITC objects and the second was the leaf-off amplitude from the 20th percentile.</p> <p>The ITC space metric was calculated as average shortest distances between object centroids per plot. These ITC objects better corresponded with the larger trees within the plot, and thus underestimated actual tree NN spacing. The field metric correlated with average crown area (section 5.2.1). Average ITC area and ITC space are linked, where an increase in one is reflected directly in the other. As mentioned before estimates of crown area can be linked to internal forest structure (Hyypä et al., 2001).</p> <p>The second model input, amplitude from a relatively low part of the forest strata is likely related to tree species (Alexander et al., 2010; Kim et al., 2009a). Given the leaf-off conditions the amplitude value could discriminate between bark and pine-needles, as in Reitberger et al. (2008).</p>
<i>Standard deviation of tree NN spacing</i>	FW LiDAR	<p>This predictive model utilises three inputs: (i) ITC average crown area; (ii) leaf-on mean height; and (iii) leaf-off echo-width kurtosis.</p> <p>This metric has relationships with crown area, as identified in section 5.2.2. As with other studies, crown area combined with height distribution metrics are a powerful explanatory metric to use (Hyypä et al., 2001; Person et al., 2002; Popescu et al., 2004; Maltamo et al., 2004).</p> <p>The echo-width metric is related to the surface roughness and slope (Lin and Mills, 2010), and thus may be related to the spatial arrangements of vegetation components within the vertical stratum.</p>
<i>Mean DBH</i>	DR LiDAR	<p>This model uses two ITC inputs, the average ITC space and mean canopy volume.</p> <p>As identified in section 5.2.2, mean DBH has a relationship with average crown area and crown base height. The ITC space and canopy volume have close relationships with these metrics; this is supported by other studies (Hyypä et al., 2001; Person et al., 2002; Popescu et al., 2004; Maltamo et al., 2004).</p>
<i>Standard deviation of DBH</i>	FW LiDAR	<p>The model uses three inputs, the leaf-off amplitude from the 65th percentile, ITC mean crown area and the leaf-on echo-width from the 25th percentile.</p> <p>The metric relating to mid-canopy amplitude may perhaps be an indication of broad species types as in Reitberger et al. (2008), in particular the presence/absence and spatial arrangements within this part of the vertical stratum (Kim et al., 2009a). As identified in section 5.2.2, the standard deviation of DBH is linked to average tree crown area. Whereas low-canopy echo-width may be related to the surface roughness and slope (Lin and Mills, 2010), and thus may be related to the spatial arrangements of vegetation components.</p>
<i>Basal area within 30x30m</i>	DR LiDAR	<p>This model utilised two inputs, these were the leaf on intensity from the 50th percentile, and the spacing of ITC objects.</p> <p>As in Brandtberg et al. (2003), different tree species can be separated by LiDAR intensity metrics, which at this level in the plot's vertical stratum may indicate some of the larger components (Moffiet et al, 2005; Kim et al., 2009a) which in turn will influence basal area.</p> <p>ITC NN space increase proportionally with average crown size, which in turn increased with total crown area, which is identified in a correlation in section 5.2.1.</p>

Table 10.1 (continued)

Field metric	Dataset	Remote sensing metrics description
<i>Percentage big trees (DBH 40>x<80cm)</i>	FW LiDAR	<p>This model utilises three inputs, these were the leaf-on kurtosis of ground amplitude, the leaf-off median non-ground height and ITC space.</p> <p>This particular metric had a number of correlations to understorey cover components, as seen in section 5.2.1, which could explain the initial ground metric. As with other studies crown area combined with height distribution metrics are a powerful explanatory metric, and have been used to predict DBH (Hyypä et al., 2001; Person et al., 2002; Popescu et al., 2004; Maltamo et al., 2004).</p>
<i>Percentage very big trees (DBH >80cm)</i>	DR LiDAR	<p>This model produced high RMSE values. It does however require one predictive input, the number of ITC objects.</p> <p>Larger trees (DBH>80cm) were present generally in lower stem densities, which is reflected in the ITC count. Even though the ITC count best corresponded with overstorey trees only, the larger trees are the ones which would make it to this level.</p>
<i>Percentage big trees (DBH >40cm)</i>	FW LiDAR	<p>This model used two inputs: the leaf-on deviation of point height and ITC NN space.</p> <p>As with other studies crown area combined with height distribution metrics are a powerful explanatory metric, and have been used to predict DBH (Hyypä et al., 2001; Person et al., 2002; Popescu et al., 2004; Maltamo et al., 2004).</p>
<i>Shannon index (by stem count)</i>	Hyper. & DR LiDAR	<p>This model for species diversity utilised four inputs, these were the red green ratio index (RGRI) for leaf-on and leaf-off hyperspectral data, the leaf-off skewness of LiDAR heights and the skewness of intensity values.</p> <p>The RGRI is related to light-use efficiency within the canopy (Gspaltl et al., 2013). The difference between leaf-on and leaf-off data may provide an indication of differences between species spectrally. With the addition of height and intensity LiDAR metrics, we can assume differences in physical structure between species may be picked up (Kim et al., 2009a; Moffiet et al., 2005).</p>
<i>Simpson index (by stem count)</i>	Hyper.	Calculated directly from ITC objects.
<i>Average height to the first live branch</i>	FW LiDAR	<p>This model had four inputs, these were: the leaf-off maximum gap within the canopy metric, average ITC NN spacing, ITC average crown volume, and leaf-on variance of non-ground amplitude.</p> <p>The first of these metrics is directly relevant, and relates to the results of a vertical profile, identifying the average largest ‘empty’ height. This is relevant due to woodland structure, i.e. overstorey, then a gap and then understorey. As identified in sections 2.2.1 and 2.2.2 there are relationships between crown area with average height to the first live branch, it is therefore reasonable to assume the remote sensing metrics are linked to ITC crown area and volume. The non-ground amplitude metric is potentially an inherent link between the differences of broad species types and their variation in canopy architecture (Andersen et al., 2006).</p>
<i>Standard deviation of height to the first live branch</i>	DR LiDAR	<p>This model utilised three inputs, these were: the maximum LiDAR derived height, the standard deviation in tree spacing and the total tree crown area.</p> <p>As a function of height, the larger the maximum height within a plot the greater the probability of disparity in stem heights per area (Matthews and Mackie, 2006). As indicated in section 2.2.2 there is a relationship between the standard deviation of the height to the first live branch and the horizontal crown area. Average ITC spacing increases proportionally with ITC area, and so it is reasonable to assume the standard deviation of crown area does the same.</p>

Table 10.1 (continued)

Field metric	Dataset	Remote sensing metrics description
<i>Average crown horizontal area</i>	DR LiDAR	<p>This model utilised three inputs, these were: the leaf-off LiDAR non-ground intensity value distribution, the standard deviation in tree spacing and the total tree crown area.</p> <p>The distribution of intensity values may be indicative of broad species type (Brandtberg et al., 2003; Moffiet et al., 2005). The average ITC spacing increases proportionally with ITC area, and it is reasonable to assume the standard deviation of crown area does the same. The values of the total ITC horizontal area metric are generally proportional to average tree area, but this accounts for a small part of the model. Both of the latter metrics corresponded well with the field measurements in their own right.</p>
<i>Standard deviation of crown horizontal area</i>	FW LiDAR	<p>This predictive model utilised one input the leaf-off amplitude of the 85th percentile.</p> <p>The relationship could be related to broad species type, as in Reitberger et al., (2008) for leaf-off data, referring to the difference between the amplitude values associated with interacting with bark and pine-needle surfaces. Kim et al. (2009a) states that intensity/amplitude values were related not only to the reflective properties of the vegetation but also the arrangement of foliage within crowns.</p>
<i>Total Crown horizontal area</i>	Hyper. & DR LiDAR	<p>This combined model utilised two inputs, these were: the leaf-on Anthocyanin reflectance index (ARI) and the leaf-on LiDAR intensity variance.</p> <p>The ARI is commonly used for the assessment of stress related leaf-pigments. Stress-related pigments include carotenoids and anthocyanins. Anthocyanins are present in higher concentrations of vegetation (Gitelson et al., 2001), supporting the assumption that crown areas are smaller in areas of higher stem density. The variance of the leaf-on LiDAR intensity may be indicative of broad species type, as in Andersen et al. (2006).</p>
<i>Standing deadwood volume</i>	FW LiDAR	<p>This model utilised three FW inputs, these were: the leaf-off skewness of non-ground amplitude, the standard deviation of ITC space, and the leaf-on echo-width at the 25th percentile.</p> <p>As with DR LiDAR, intensity or amplitude is linked to total deadwood biomass (Kim et al., 2009b; Martinuzzi et al., 2009). According to section 5.2.1 standing deadwood volume is correlated with the standard deviation of tree crown area. The average ITC spacing increases proportionally with ITC area, it is reasonable to assume the standard deviation of crown area does the same. For the final metric, Mücke et al. (2012) states that the FW return echo-width relates to small height variations of scattering elements within the footprint of the laser beam, and was considered a means of inferring surface roughness. Forest ground-level and downed stems were assumed to have smooth surfaces, whereas other vegetated elements, such as shrub vegetation, were considered to be rougher. Standing deadwood is not on the forest floor, but the principle should still apply.</p>
<i>Standing deadwood decay class</i>	FW LiDAR	<p>The estimation of the standing deadwood decay class was best predicted using two inputs, these were: the leaf-on amplitude from the 35th percentile and the leaf-on number of returns from the 20th percentile.</p> <p>As with DR LiDAR, intensity or amplitude is linked to total deadwood biomass (Kim et al., 2009b; Martinuzzi et al., 2009), and it appears deadwood state. The 35th and 20th percentiles which on average relate to approximately 4-7m are where the surveyor would expect to find the deadwood items.</p>

Table 10.1 (continued)

Field metric	Dataset	Remote sensing metrics description
<i>Downed deadwood volume</i>	FW LiDAR	<p>This model utilises two inputs, these are: the leaf-off skewness of non-ground amplitude and the leaf-on deviation in heights.</p> <p>As with DR LiDAR, intensity or amplitude is linked to total deadwood biomass (Pesonen et al., 2008), although these measurements are not from the ground-level, there would seem to be a link to overstorey structure, especially when considering the second metric included within this model's construction.</p>
<i>Downed dead wood decay class</i>	Hyper.	<p>Only one input to the regression model is included, this is the simple ratio index (SRI) derived from leaf-on data.</p> <p>The SRI is a broadband greenness VI. As mentioned in Pu et al. (2008) there is a possible link between vegetative greenness and mortality. Alternatively, deadwood volume was generally higher in deciduous areas, which have different index values and ranges in comparison to coniferous areas.</p>
<i>No. saplings</i>	DR LiDAR	<p>The regression model employed three inputs; these were the leaf-on 60th percentile for intensity values, 60th percentile for number of returns, and the mean intensity value of all points.</p> <p>The number of saplings appears to be related to tree spacing, crown area and overstorey species number, according to sections 5.2.1 and 5.2.2. Here the assumption is made that LiDAR measures here are related to both species and vertical structure (Brandtberg et al., 2003; Moffiet et al., 2005).</p>
<i>No. native saplings</i>	Hyper. and DR LiDAR	<p>This particular model employs two inputs, these are: the modified red edge simple ratio index (MRESRI) derived from leaf-on hyperspectral data and the skewness of the leaf-off LiDAR intensity values from the ground-level.</p> <p>The MRESRI is a narrowband greenness VI and has relationships with the amount of photosynthetic material and canopy architecture (ENVI-online-help, 2005). Leaf-on data would also highlight the difference between coniferous and deciduous species types. Ground-level intensity distribution metrics may be related to different species types (Brandtberg et al., 2003).</p>
<i>No. of sapling species</i>	Hyper.	<p>This predictive model utilises three inputs, these are: the leaf-on Anthocyanin reflectance index (ARI), the normalised differenced vegetation index (NDVI) and the number of classified ITC objects of a native tree species.</p> <p>The ARI is commonly used for the assessment of stress related leaf-pigments. Anthocyanins are present in higher concentrations of vegetation (Gitelson et al., 2001). As identified in Sims and Gamon (2002) greenness indices are related to canopy leaf-area and architecture. The final input, the number of ITCs of a native species accounts for a small amount in the model. Thus, overstorey composition and structure are determinants to sapling composition.</p>
<i>No. seedlings</i>	Hyper. and FW LiDAR	<p>This predictive model utilises two metrics, the first is a measure of the variance in the distribution of leaf-off heights, and the other is an estimate of overstorey species number from classified ITC objects.</p> <p>According to section 5.2.2 there are partial relationships between the number of seedlings and standing deadwood decay level, which in turn is correlated with crown area. We must assume that the LiDAR metric is related to vertical structure. The relevance of the estimate of overstorey tree species number here must be related to the weak proportional relationship where a rise in one sees a rise in the other.</p>

Table 10.1 (continued)

Field metric	Dataset	Remote sensing metrics description
<i>No. native seedlings</i>	DR	<p>This particular model utilises two inputs, these are the standard deviation of ITC spacing and the ITC crown volume.</p> <p>According to sections 5.2.1 and 5.2.2 the number of native seedlings is related to the standard deviation of plot DBH. Thus, by extension the ITC attributes which can be used to estimate DBH can be used to estimate the number of seedlings of native species (Hyypä et al., 2001; Person et al., 2002; Popescu et al., 2004; Maltamo et al., 2004).</p>
<i>No. of seedling species</i>	Hyper and FW LiDAR	<p>This regression model requires the following inputs, the simple ratio index (SRI) as calculated from leaf-on hyperspectral data and the standard deviation echo-width values from the ground classified FW LiDAR.</p> <p>The SRI is a broadband greenness VI related to the forest woodland canopy architecture and the penetration of radiation into the canopy according to Chen, (1996). Mücke et al. (2012) states that the FW return echo-width relates to small height variations of scattering elements within the footprint of the laser beam, and was considered a means of inferring surface roughness. Forest ground-level and downed stems were assumed to have smooth surfaces, whereas other vegetated elements, such as shrub vegetation, were considered to be rougher.</p>
<i>Seedlings Shannon Index for native species</i>	Hyper and FW LiDAR	<p>The predictive model required two inputs, the first was a leaf-on LiDAR metric showing the absolute deviation of height points, and the second was the classified ITC species count. The error for this metric was relatively high concerning the index range (RMSE 0.70).</p> <p>According to sections 5.2.1 and 5.2.2 the seedling Shannon index has relationships with the number of tree species within the plot. The assumption is that of the FWs absolute deviation of height values is related to these forest structural metrics, whilst the ITC derived species count is a reasonable approximation of actual tree species number.</p>
<i>No. Vascular Species</i>	FW LiDAR	<p>This model is predicted by one FW LiDAR metric input, the leaf-on variance of non-ground height values.</p> <p>The assumption is that the FW variance of non-ground height values is related to forest structural metrics, and thus the diversity of vascular species.</p>
<i>The percentage of non-vegetated cover or bare soil</i>	FW	<p>This predictive model required inputs from three metrics, these were: the number of leaf-on returns from the 20th percentile, the estimate of canopy depth from leaf-off data and the standard deviation of ITC spacing.</p> <p>The 20th percentile corresponds to approximately 4-6m, dependent on canopy height. One must make the assumption that more returns at this vertical level are indicative of a more developed understorey. Likewise, the estimates of canopy depth may indicate greater or less vegetation cover, for example in considering the amount of light penetration (Jensen et al., 2008). The field-level metric of the percentage of non-vegetated cover is correlated with the standard deviation of DBH and basal area within the plot as in section 5.2.1. As seen with other predictions ITC space is related to crown area and this is related to DBH as in Hyypä et al. (2001). It should be noted that this metric was estimated visually within the field.</p>

Table 10.1 (continued)

Field metric	Dataset	Remote sensing metrics description
<i>Sum of diameter (DBH) differences (cm)</i>	FW	<p>There were four inputs to the predictive model; these were (i) the leaf-off absolute deviation of the mean echo-width, (ii) the leaf-off amplitude recorded at the 90th percentile, (iii) the leaf-on number of points from the 20th percentile, and (iv) the leaf-on kurtosis of echo-width from non-ground points.</p> <p>Echo-width may be related to the surface roughness and slope (Lin and Mills, 2010), and thus may be related to the spatial arrangements of vegetation components. The combination of height, amplitude and echo-width metrics may be an indication of broad species type and woodland structure (Reitberger et al., 2008; Heinzel and Koch, 2011).</p>
<i>Index of vertical separation (CSDI#2)</i>	FW	<p>This model utilised two inputs, these were the leaf-on skewness of non-ground height values and the leaf-off amplitude at the 85th percentile.</p> <p>This index is based upon the difference between the three largest and three smallest DBH values in the plot. The first metric is likely to be related to vertical structure, whereas the leaf-off amplitude relates to the difference between pine needles and bark (Andersen et al., 2006; Reitberger et al., 2008).</p>

10.5 A review of remote sensing derived condition indices

There are very few examples within the surrounding literature of estimating forest condition status from remote sensing data. As previously mentioned, Simonson et al. (2013) outlines an approach to combine LiDAR and multispectral data, in addition to spatial pattern metrics to classify various habitat and land cover classes and ultimately a gradient of condition status. Other authors, such as Borreet al. (2011), summarise the needs of the Natura 2000 monitoring approach and the potential benefits of applying remote sensing data and analysis techniques.

Once the regression models were completed and validated it was possible to map the field metrics, and to apply conventional condition index assessment methods to the dataset. The concept of the condition index is simply to generate a summary value which incorporates the condition considerations of both forest structural diversity and the composition of the biotic and abiotic components of the ecosystem. Within the course of this study six established condition index calculations were applied to the metrics extracted from the remote sensing datasets. Of the six condition indices, two were calculated directly from DR LiDAR derived ITC information, three utilised only area-based metrics as derived through statistical approaches, and one combined inputs from both ITC and area-based metrics. The results of this will be discussed in the following sections. Comparable forest condition assessment approaches as identified in the surrounding literature are also discussed and contrasted and the best condition assessment method identified.

10.5.1 Condition assessment methods tested

The vertical evenness index was computed through the use of ITC inputs in order to gauge the distribution of canopy cover vertically through the forest plot. The method produced results which corresponded well to field data in areas where canopy openness was high – i.e. each of the layers within the plot stratum could be observed. However as canopy openness decreased, and tree crowns merged into one another, the accuracies of the index fell. This was especially prevalent in deciduous woodlands, this again highlights the limitations of the ITC method employed within this research project, where total tree counts were underestimated and the presence of suppressed trees were missed. In order to solve this issue a different approach to identifying forest horizontal cover at different vertical levels, rather than from the CHM alone, could potentially use height metrics extracted from the point cloud, such as in Lesak et al. (2011), or by utilising statistical approaches.

The Clark-Evans aggregation index was computed using the distance metrics calculated in-between ITC centroid points per plot. The ITC derived value always produced index values above 1 (which indicates a random horizontal stem distribution), while this was often similar to many of the field plots; it did not reflect those with index values below 1. Upon mapping this index value it was clear that only a limited difference in structure could be inferred, and was generally a poor indication of true forest structure. Again, this highlights the limitations of the ITC delineation approach employed in this project for the detection of tree numbers and locations. RMSE estimates of tree NN spacing calculated from ITC data was 4.2m, indicating a substantial source of error. As with the above, a method able to delineate the position of suppressed trees and cope better with the crowns of deciduous trees with more non-regular forms, (Leckie et al., 2003a; 2003b), would potentially provide a better means of calculating this index.

The diameter differentiation index was calculated using two area-based metrics as predicted statistically from remote sensing metrics. These metrics were an estimate of tree density and the sum of the diameter differences within the plot. The latter was predicted through a regression model, with an FW LiDAR model producing a RMSE of 19.5cm. The estimated index values performed well in comparison to those recorded in the field for all but three of the 20 validation plots. The index value ranged from 0-1, the overall RMSE was calculated to be 0.19. The mapping produced using this index showed a correspondence of values to the different FC compartment boundaries, and therefore forest structural types. The deciduous

semi-ancient woodland contained the largest amount of variation in index value. The majority of the predicted map grid-cells occupied the range of 0.8 to 1.0, and overall did not show much difference between values within deciduous and coniferous woodland or between compartments. Thus this method was considered poor.

The complexity index (HC) was calculated using four area-based metrics which were statistically derived from the remote sensing datasets. This particular index typically requires a target value in order for assessments to be made, whereas here a larger value was considered to indicate a better condition. Index results for this approach proved to be highly variable. Plot 30 again caused issues and was greatly underestimated, increasing overall model RMSE. The remainder were much better predicted and were within ± 50 of the field index value i.e. within 20% accuracy. The resulting map again corresponds well with FC compartment boundaries and structural types. The HC index is more indicative of tree size parameters. There are also obvious differences in deciduous and coniferous area values; however the index tended to skew towards high or low values for each of the woodland types respectively.

The complex stand diversity index (CSDI) is composed of four component indices many of which require the explicit measurement of individual tree metrics. Approximately 20 input metrics are required for this computationally complex index, which incorporated metrics related to species proportion, and the differentiation of tree stem spacing, DBH and crown dimensions. The field index range of values was 4 to 13. Overall index values from the field corresponded well with those derived through remote sensing, producing a RMSE of 3.5. When the index values were mapped across the study site, again the values corresponded well with compartment areas and structural types. The deciduous area exhibited the most variability. The CSDI relates more to structural diversity which can transcend both the broad types of deciduous or coniferous woodland, however higher values were typically located in deciduous areas.

The CSDI utilised various inputs from ITC information, for example, the index relating to the estimation of species proportions, which resulted in the poorest accuracy of the four input component indices. Again this is because of the ITC underestimation of tree numbers and the missing of suppressed trees. There were also concerns about the accuracy of spacing estimates made from ITC positions.

ITC estimates of the three smallest and three largest crown diameters produced relatively high correspondence to field data values, as did the three largest and three smallest stem NN distance measurements once appropriate corrections had been made. The component index for characterising the differences between the three largest and three smallest DBH values was estimated statistically using regression, resulting in a RMSE of 0.13, as derived by FW LiDAR.

The score-based index employed 17 input metrics covering compositional and structural metrics for both overstorey and understorey components. Each of the inputs scored a value of one if they exceeded a target value, whereas they scored a zero if they did not. A 'perfect' site would therefore receive a value of 17. Overall remote sensing derived estimates corresponded well with field data calculated values, producing a RMSE of 2.5. The mapped index values corresponded well to the different FC compartment boundaries and structural types. It should be noted however that the targets which the indicators had to reach were defined from deciduous ancient and semi-ancient forest examples within the literature in Cantarello and Newton (2008). This may have the consequence of biasing the index to deciduous areas. It should be noted that some of these targets, especially for understorey metrics, were high and thus were not achieved within the New Forest study site, for example the number of seedlings.

The VE index, Clark-Evans aggregation and the diameter difference indices proved to be insufficient for mapping condition due to the involved error. Whilst each of the remaining three condition indices had merit, it was concluded that the score-based index provided the most comprehensive indication of woodland condition, bearing in mind sources of error and its ability to discriminate between the different forest structural types when considering map products. Its use of area-based metrics only reduced a number of uncertainties about ITC related measurements.

Initially there were concerns that adding different remote sensing estimated metrics together in the course of calculating different condition indices would increase error overall and lead to poor index estimates. This error is apparent from the results presented in Appendix F section F.2, especially when combining model results developed from different remote

sensing datasets. It should be noted however than in all cases the additional uncertainties were small. The results presented here indicate that, while there is error present, it is possible to produce estimates of forest condition via conventional index methods to a reasonable level of accuracy.

10.5.2 A comparison of the forest condition index methods

Each of the six condition indices were mapped across the entire study site, and the individual 30x30m grid-cell values were categorised from 1-5 in order to make them directly comparable with one another. The VE, Clark-Evans aggregation and diameter difference indices were removed from the comparison owing to problems related to ITC metric underestimation of stems or index values which were too similar. In the maps created from the remaining condition indices, HC, CSDI and score-based, a number of spatial patterns could be observed which corresponded with known FC forest compartment boundaries. A mean and standard deviation of the three categorised index values was computed for each mapped grid-cell in order to assess the similarities or dissimilarities of the three remaining indices.

There was a high degree of correspondence between the output maps from the HC, CSDI and score-based indices. High index values for all three were almost exclusively found in deciduous dominated woodland, with only a limited number of high index values located within coniferous dominated woodland. Consistently higher scoring deciduous woodland is typically dominated by beech and oak species from areas initially planted around 1800. Even when combined with trees planted more recently, high index values were still achieved. These areas are defined by the presence of large trees (i.e. DBH values) with variable sizes, with higher numbers of species and the presence of deadwood. These areas may also contain understorey components which are lacking in coniferous areas. The highest values in coniferous areas were for compartments with a mix of species types. These areas were also defined by the presence of large trees with variable sizes, higher numbers of species and the presence of deadwood. The majority of coniferous areas however exhibited relatively small tree sizes (often with little variation in their respective sizes) few species with little or no deadwood or understorey.

There are a number of examples where there were dissimilarities between the three indices; this was encountered primarily within deciduous woodland with long continuity (i.e. planting dates of ~1800). The HC index tends towards an almost binary output depicting coniferous or deciduous woodland types. The HC inputs are weighted more towards tree size which is most typical of deciduous woodland. The CSDI on the other hand, is more related to species and structural diversity, which can be present in both deciduous and coniferous woodland, although the presence of the latter is limited to few areas. There were a number of examples where the CSDI overestimated the index value in grid-cells where large canopy gaps occur, or failed in other areas due to ITC metric issues. The score-based index combined both tree-size and diversity metrics at a relatively basic level in addition to many other factors, such as deadwood volume and understorey populations. The score-based index also did not directly use ITC-metrics and thus should be less influenced by its associated error.

The primary driving factors of woodland condition would seem to be related to the average size of the trees (i.e. DBH) and the variability of the sizes over the plot which is especially a feature of oak and beech woodland with long continuity, for example with planting dates approaching 1800. Given the correct conditions, the much younger coniferous compartments can reach similar high index values, but this is rarely the case. The presence and volume of deadwood seems to follow these patterns. However, understorey metrics such as vascular species number, and the number and composition of seedlings and saplings seem not to. Owing to the limited levels of understorey encountered within the New Forest study site, finding a link may not be possible.

Thus, the score-based index combines elements from the HC and CSDI indices and important features such as understorey and deadwood, which may not be linked to tree size and diversity due to factors such as forest management. In addition, it avoids some of the issues inherent in using ITC metrics directly and, overall provides the most representative condition assessment method.

10.5.3 Assessing condition in the New Forest study site

To reiterate, the condition index calculation method was developed by Cantarello and Newton (2008). A remote sensing derived condition indicator map was produced covering an area of approximately 19km². Each of the 30x30m cells were assessed against the desired

condition level set in the aforementioned study. The produced condition indicators were then assessed for forest management purposes.

When interrogating the condition index map there are obvious differences between coniferous and deciduous woodland areas, or, more specifically, between woodland compartments, i.e. indicative of different compositional and structural attributes. Figure 10.1 gives an overview of the produced condition index map for the whole study site. A number of woodland compartments also exhibit condition gradients within their own boundaries. The index is an indicator to guide assessment of the condition of the area in question, as some sites may be sufficient in some aspects of condition and deficient in others. The maximum index value recorded was 14, so none of the sites met all the targets set for the condition assessment.

Seven example woodland areas will be assessed here, these are: Frame, Tantany and Denny Woods for semi-ancient unenclosed woodland, Hawkhill, Frame-Heath, Stockley and Denny-Lodge inclosures for managed and enclosed plantation woodlands.

The first of these is Tantany Wood which is semi-ancient deciduous woodland in the east of the study site. In the west of Tantany Wood there are consistently higher values. These areas contain the largest and oldest deciduous trees, which were initially planted in 1800. There are lower and more variable index values to the east in more recently planted areas (1930), dominated by birch, and the areas which boarder the heathlands to the north. This general pattern is reflected in many of the 17 input metrics. The older planted woodland to the west has variable low to high stem densities, typically a high SH index value and high and variable average DBH values. Total basal area level is moderate, but is higher in the recently planted areas. More recently planted areas are more varied in species number. Standing deadwood volume is higher to the west but generally has a lower decay class. Downed deadwood volume is again higher to the west, but also has a higher decay class than to the east. The same is true for seedling number and the number of vascular species. This area of Tantany Wood contains some of the highest numbers of seedlings within the study site. Sapling number is more consistent apart from smaller values in recently planted areas. This is likely consistent with the high number of holly trees existing in the understorey here.

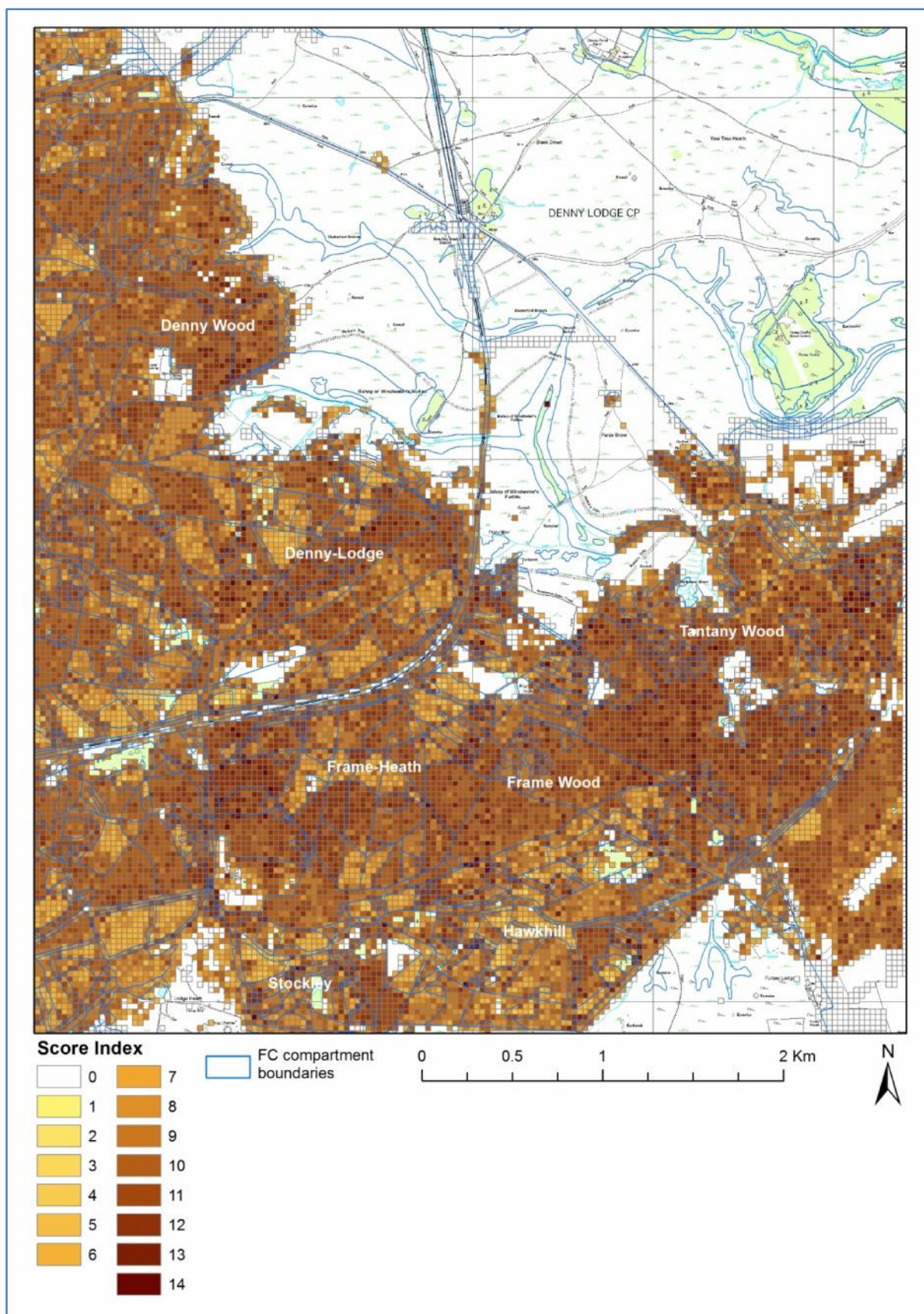


Figure 10.1 – The score-based index applied across the whole study site. Base Map layer is © Crown Copyright/database right 2010. An Ordnance Survey/EDINA supplied service.

Frame Wood borders Tantanty Wood to the east, where there is an obvious decline in index value. Again the planting date was 1800. The decomposed index attributes correspond to a generally low stem density, but with a relatively high SH index. Mean DBH is high, and varies within each plot. Total basal area is low. Standing deadwood volume exists in moderate to high levels, as does decay class. Downed deadwood volume exists in high level; however the decay class varies from moderate in the east and high in the west. The number of saplings and seedlings overall is low, likewise the number of vascular species is low overall per grid cell.

Denny Wood is located to the north-west of the study site and exhibits slightly lower index values for deciduous woodlands. The total number of trees is low, giving way to relatively open areas throughout. High index values are recorded for clusters of larger trees which exist in this extent. Mean DBH value is highly variable, as is total basal area. Standing deadwood volume is variable, whereas decay class is high. Downed deadwood volume is moderate across the woodland, with higher values to the east. Sapling number is low; however there are pockets of higher numbers. Seedling number increased from low to moderate from east to west. The number of vascular species is low, with very high numbers occurring within some clearings.

The enclosed plantation woodlands typically contain 10 or more compartments, which can contain coniferous, deciduous and mixed woodland types. The deciduous, and mixed woodland to a limited extent, follow the same pattern as the above, however coniferous woodlands, all of which were planted from 1940-1980, have consistently lower condition index values.

Hawkhill inclosure lies to the south of the study site. Within the coniferous woodlands there are differences in index values between compartments, this is typically because of variation in average height and/or stem density. Stockely inclosure has a relatively small extent, and shares a border with Hawkhill to the east. The inclosure contains many small compartments in close proximity which contain many different species. This resulted in a mix of index values, however deciduous clusters scored higher. The Frame-Heath inclosure lies at the centre of the study site and again contains a large mix of woodland types. There are differences between deciduous compartments in the west. The largest values are for plots with two planting dates, 1852 and 1958, resulting in higher stem density and basal area

values. There are lower values in proximity to the farm located in the south-west. The proximity to the rail line does not seem to have an effect on the vegetation metrics. The Denny-Lodge inclosure is located to the west of the centre of the study site. Again there is an obvious distinction between deciduous and coniferous dominated woodlands. Different values are present in mixed conifer and deciduous woodland.

Conifer compartments typically have little to no understorey, but there are some exceptions to this, such as areas in Denny-Lodge where there was a strong understorey component, which was confirmed through fieldwork. There was little or no deadwood, relatively low DBH values, with negligible differences in DBH value and low species diversity (SH index) within a plot. This means that the grid-cells which intersect with these areas do not score very highly for anything except stem density and basal area. Plantation forests such as these are managed for productivity at the expense of other objectives. Deadwood for example is periodically cleared and left in scattered piles throughout the compartments, which was observed during fieldwork. This results in a homogeneous composition and structure within a compartment.

Index value, especially in deciduous woodland is higher the further away from public access points, car-parks and caravan sites. This is probably because of the fact that these access points are on the boundary between the forested area and heathland or grassland, rather than disturbance, although this is still possible.

Overall, a number of trends are evident which are of importance to forest managers. The amount of regenerating trees, i.e. saplings and seedlings, is low across the study site, especially for coniferous woodlands, thus the target value for the score index is not met for much of the study site, a trend recognised by Mountford et al. (1999); this is because of grazing of a high number of large ungulate species, such as deer, which are present throughout the forest. When a compartment has been cleared and replanting operations have begun, the area is fenced off to stop ungulate browsing and allow the trees to grow.

The number of vascular plant species is also typically low apart from within a small number of coniferous compartments within Frame-Heath and Denny-Lodge, which corresponds with relatively high canopy openness. The occurrence of both standing and downed deadwood seems somewhat limited to deciduous woodland types, although Japanese and hybrid larch compartments also exhibit high downed deadwood volume amounts. The movement of

deadwood will influence nutrient cycling and affect the availability of resources for both flora and fauna. In terms of forest structure and composition, it is obvious that many of the compartments have very similar structure internally which is related to the age structure caused by distinct planting dates (Putman, 2010).

It is clear that the current forest management has had a great deal of impact on the structural and compositional components of the forest, causing similar attributes within each compartment owing to distinct planting dates, understorey regeneration and diversity being negated by no ungulate control, and deadwood resources are removed in plantation areas. Each of these elements has an impact on overall condition, thus the above analysis highlights some of the shortcomings of current management approaches.

10.6 Implications and future work

The methods developed in this research can be used to support both forestry and conservation assessments at scales from the field-plot (30x30m) to the landscape scale, although further refinement is required for the successful retrieval of relevant metrics in some very dense forest structural types. However this study has provided a solution, while partially incomplete, to the consistent retrieval of 35 metrics which can be used to assess forest condition. Four metrics could either not be modelled or provided poor estimates (i.e. NRMSE > 0.4). The method presented here has set a baseline against which the success of future work can be assessed and has provided an insight into some of the limitations of such an approach and the obstacles that need to be overcome.

This section discusses the relevance of the current research in relation to the wider context of application of airborne remote sensing analysis for forest management activities for both forestry and conservation activities. A description of a number of potential improvements and future work follow this.

10.6.1 Usefulness in the wider field

The methods and analysis presented within this study demonstrate the processing and use of data from the relatively new FW LiDAR technology. The application of FW LiDAR in an operational sense is limited in the surrounding literature. The processing techniques are still in development, and may potentially offer additional useful techniques and benefits to analysis in the future.

The methods outlined in this thesis have potential applications in the areas of forest management inventory and conservation. The approaches reported here demonstrate the comparative merits of hyperspectral, DR and FW LiDAR data for the prediction of various conventional and unconventional metrics within a forest context. The ability for a large amount of ‘real world’ forest metrics to be estimated from remote sensing datasets in isolation and in combination was demonstrated.

10.6.1.1 Ecology and habitat quality assessment

The assessment of habitat quality is an important issue globally. Many European countries are required to assess the quality of habitats within their borders, rather than simply identify a given vegetation patch as a habitat. An example assessment initiative is Natura 2000.

Forest conservation and management are important and complex processes, which have significant implications both environmentally and economically. An efficient ecological assessment approach requires detailed knowledge of species composition, distributions and structure. The conventional approach is to record these characteristics manually through field survey and extrapolate the results over large areas. Various forest condition assessment criteria were found, and the estimation of many of these metrics was tested through the course of this research project, including the typical means of depicting a site’s overall condition. The six indices tested here utilised various estimations of field metrics relating to vertical canopy cover diversity, tree DBH diversity, tree crown size diversity, species diversity and spatial separation; many of which could be estimated with relatively low error values. The best assessment method utilised a total of 17 overstorey and understorey, structural and composition metrics which are required in conservation status assessment, as presented in Cantarello and Newton (2008). Its implementation here produced estimates which closely corresponded with field data, and which allowed condition to be assessed continuously throughout a large study site region.

10.6.1.2 Habitat suitability modelling

The approach defined here could also have applications in predicting habitat types indirectly through proxies derived from remote sensing measurement. Spatially explicit data about the 3D structure of forests is required for the assessment of biodiversity and wildlife habitat, and are important factors explaining (i) the presence of many wildlife species, (ii) the functional use of habitat (e.g. nesting), and (iii) the overall diversity of wildlife species in forests (MacArthur and MacArthur, 1961). A number of authors have utilised the 3D information from LiDAR data to estimate bird habitat (Lesak et al., 2011; Martinuzzi et al., 2009). In particular the presence of deadwood and understorey vegetation was directly related to the presence of a number of avian species in the study put forward by Martinuzzi et al (2009). Likewise McDermid et al. (2009) assessed the likelihood of the presence of grizzly bears by assessing forest elements such as canopy closure, dominant vegetation species, and vegetative moisture content.

The mapping of the presence of forest attributes necessary for the presence of certain flora and faunal species is possible utilising the techniques described within this study. Forest compositional and structural components can be mapped spatially across wide areas, as demonstrated in Chapter 9.1. Relevant calculations could then be applied to predict the habitat probability of certain species presence in a similar manner to a forest condition index.

10.6.1.3 Forestry

One of the areas where the approach demonstrated in this study has potential is that of precision forestry. As outlined in Matthews and Mackie (2006) there is a requirement for knowledge of how many trees exist, what species they are and their relative sizes for a defined area, in order to make accurate predictions of timber volume yield and ultimately the commercial value of the timber therein. These requirements can be fulfilled using the approach demonstrated here, where estimates of tree number, species type, average DBH size, variation in DBH size, and basal area can be predicted within a reasonable accuracy level (NRMSE <0.4). While timber volume was not estimated within the scope of this thesis, it has been estimated successfully in a number of other studies which utilise similar regression-based predictive approaches, such as Naesset (2002) and Persson et al. (2002). It is also worth noting that similar approaches have already seen operational deployment in Nordic countries for forest resources inventory (Naesset et al., 2004). The approach defined here has been successfully applied to a range of different forest types. With minor

modifications this approach here would also allow carbon stock assessment. The national reporting of carbon sources and sinks is required to fulfil national obligations related to climate change (Rosenqvist et al., 2003).

10.6.1.4 Further uses with minor modifications

If appropriate time series data are acquired, there is the possibility of comparison of changes over time for large areas, are at a much finer resolution than possible through fieldwork alone. For example, Falkowski et al. (2009a) demonstrated that it was possible to characterise forest succession through statistical methods when applied to LiDAR data. Whereas Yu et al. (2006) could estimate tree growth increments between different LiDAR data acquisitions five years apart. For each data acquisition, similar methods to those applied here could be implemented and metrics estimated. The metrics for each date could then be compared to determine differences and similarities used to forecast potential future differences. Adding a temporal dimension to this kind of analysis would add the benefits of being able to assess changes over time, such as growth, productivity and damage more accurately and completely than through field assessment alone. This would also add the benefit of being able to assess the impacts of various management activities through time.

10.6.2 Potential improvements and future work

The methods developed in this research can be used to support forest condition assessment at the field plot-scale over wide areas. A number of refinements can potentially be made to mitigate against some of the limitations identified or provide improvements and alternative outputs.

10.6.2.1 ITC underestimation

The underestimation of tree stems by automated ITC delineation procedures requires future attention. Alternative ITC delineation techniques, which do not utilise a rasterised CHM in order to locate tree crowns, could be employed. A number of alternative ITC delineation approaches exist which operate on the point cloud rather than derived raster height models, one such approach utilises the 3D distribution of LiDAR points expressed as volumetric pixels, or voxels. Wang et al. (2008) presents an experimental method to define individual tree crowns by identifying cluster features within the voxels using airborne LiDAR data from a spruce dominated site in southern Germany. The basic concept of single tree extraction is to trace the outlines on the projection images from top to bottom using projection images at

different height levels. Using a hierarchical morphological opening and closing process, the structural elements of the tree crowns can be identified at each layer. Reitberger et al. (2009) outlines a voxel based approach utilising small-footprint FW LiDAR data to detect trees in the forest canopy through a normalised-cut 3D segmentation approach, and which functions even in the lower forest layers. The study site was in South-East Germany, and produced detection accuracies of 86% for upper-canopy trees and a total accuracy of 58% for all trees within the plot.

10.6.2.2 ITC suppressed trees

The lack of capability for detection of suppressed-trees and trees within the understorey through automated ITC delineation methods is a known issue with such techniques, as stated by Karrtinen et al. (2012), who indicated that further processing of the point cloud would be required.

The ITC counts derived in this research project do not adequately represent the total tree stem number recorded within the field plots, as in Hyyppä et al. (2001). One must assume the presence of suppressed trees or those located within the lower strata of the forest canopy cause this discrepancy. The research of Kaartinen et al. (2012) into employing different automated ITC delineation algorithms also highlighted the need for further processing to detect suppressed tree data at the point level. Lee and Lucas (2007) utilised a method using LiDAR data to derive a height-scaled crown openness index (HSCOI) raster layer in order to identify stems in both the overstorey and sub-canopy for a forest in Queensland, Australia. The HSCOI functions by representing the density of forest canopy components by quantifying the penetration of the LiDAR returns. The local minima in the raster layer correspond with the upper canopy, such as above the central tree stem and larger proportions of major branches. An automated image analysis approach was implemented to locate and map the tree stems from this layer. Height and DBH estimates could be made for these suppressed trees using an empirical relationship between tree height and the minimum HSCOI value (Lucas et al., 2006a; Lucas et al., 2008b).

An alternative approach was explored in Maltamo et al. (2005) where the existence and number of suppressed trees was examined through the use of the height distributions of the returned laser pulses. A histogram thresholding method was applied to the recorded height

distributions in order to separate different tree stories. Finally, the number and sizes of the suppressed trees were predicted using regression models. The results showed that multi-layered stand structures can be recognised and quantified using quantiles of return height distribution data. However, the accuracy of the results is dependent upon the density of the dominant tree layer.

It is also worth experimenting whether empirical relationships exist between the number of overstorey trees and suppressed trees, for example, based on height and crown extent relationships. Another alternative would be the use of a LiDAR system which acquired data at a higher sampling rate or utilised a smaller footprint size to allow greater penetration into the canopy, or employ FW LiDAR which, as demonstrated here, can resolve a higher number of points within the canopy.

10.6.2.3 Overstorey species classification

The delineation of ITC objects is of direct relevance to the classification of trees within the overstorey, as better defined tree crowns would improve classification accuracy overall. Additional spectral datasets were available for this study from the Specim Hawk dataset, which recorded shortwave infrared wavelengths. Although at a different spatial scale Lucas et al. (2006a) increased classification accuracy using shortwave infrared data. The species of a tree is only one of several factors that affect the realised shape and spectral response of an individual tree crown. Other factors such as terrain, environment, competition and genetic variation have influences upon the spectral response of the tree as well. Hill et al. (2009a) also demonstrates the potential benefits of using time-series data to better classify overstorey tree species.

The addition of LiDAR attributes may also enhance species classification (Dalponte et al., 2008; Hill and Thompson, 2005; Moffiet et al., 2005; Simonson et al., 2013). Metrics related to canopy height, texture and difference in penetration of laser pulses into the canopy of different species could be of benefit here. As stated in Vaughn et al. (2012), trees can be classified without using spectral data, using segmented 3D information from LiDAR combined with the FW LiDAR metric related to the distance between peaks in each waveform.

10.6.2.4 Understorey species classification

The detection of suppressed trees or trees in the understorey would add the issue of identification of its species. There would be the potential to utilise structural information from the LiDAR data, as in Moffiet et al. (2005), as well as known community association for different environments, for example for different soils and topography. When utilising hyperspectral data it may be possible to spectrally un-mix the information for tree crown objects (Tits et al., 2012).

10.6.2.5 Hyperspectral derived metrics

A large number of vegetation indices (VI) were calculated from the Eagle hyperspectral data for this project and were important for estimating a number of forest condition indicator metrics. Some of these spectral indices describe similar information; however it is unknown which may provide the superior products for estimating forest metrics relating to condition. For example there are broadband greenness and narrowband greenness indices, the latter of which make greater use of reflectance measurements in the red and near-infrared regions of the spectrum which allows much greater penetration depth through the canopy than red measurements employed in broadband indices (ENVI-online-help, 2005). Thus, narrowband measurements in the red edge allow these indices to be more sensitive to smaller changes in vegetation health than the broadband greenness VIs, particularly in conditions of dense vegetation where the broadband measures can saturate. Further experimentation with the VI outputs is required in order to determine the best ones.

A number of additional VIs can be computed if measurements from the shortwave infrared range are utilised, for example with the inclusion of the Hawk dataset. These VIs could be used for the quantification of biophysical values for canopy nitrogen content and dry or senescent carbon for example (Treitz and Howarth, 1999). There are also numerous other features which can be extracted from hyperspectral datasets. Additional metrics can be extracted directly from the hyperspectral data without the need to calculate VIs or dimensionality reduction, where information from individual bands (each representing a discrete range of the electromagnetic spectrum), which are linked to vegetation characteristics, can be extracted (e.g. Lucas et al. 2008a). Incorporating data from individual spectral or dimensionally reduced bands (e.g. MNF, PCA) (Hill and Thomson, 2005), or image texture features (Heinzel and Koch, 2013) may also provide useful metrics for predicting forest information. Area-based summaries of hyperspectral derived imagery such

as spectral bands or dimensionally reduced data (e.g. mean, max, min, etc.) could be incorporated as with the VIs in this project.

10.6.2.6 Additional LiDAR metrics

The number of LiDAR metrics describing the statistical distribution of height and intensity values was initially determined by the RSC LAS tools software. Thus in comparison to other studies a number of more complicated metrics could not be computed, for example canopy profile metrics (e.g. number of vegetation returns $> 1\text{m}$ and $\leq 2.5\text{m}$ in height), voxel metrics, and metrics which indicate clustering of returns, or the stratification of points by return number. Some of these metrics can be computed in other pieces of software, such as LASTools, but unfortunately the processing of data through RSC LAS tools alters the LAS file structure often making it unreadable to other software which caused further complications.

The SPDlib software was employed relatively late through the course of this project and as such could not be used for both FW and DR LiDAR datasets. The SPDlib software provided overall more potential LiDAR metric outputs and user control for calculating additional non-standard metrics, including the capability to compute HSCOI. The software also has the ability to process both DR and FW LiDAR data. It should be noted that SPDlib was still in development at the time of writing this thesis and may be able to produce more outputs in the future. Other authors have identified additional metrics to extract from the waveform data such as the backscatter cross-section and coefficient. The backscatter cross-section of the laser pulse can be used to understand the characteristics of the return signal (Alexander et al., 2010). A target's backscatter cross-section depends on its size, reflectivity of its surface, and the directionality of scattering of the pulse reflection caused by the target's geometric shape. The backscatter coefficient is a normalised measure of the backscatter cross-section irrespective of area and footprint (Wagner et al., 2008; Woodhouse, 2006).

There exists metrics which have been derived from either airborne or spaceborne large-footprint or profiling waveform LiDAR systems, such as LVIS, which could potentially be modified for experimental use with small-footprint FW LiDAR such as that utilised in this project. An example of such a metric is the height of median energy (HOME) (Anderson et al., 2008; Goetz et al., 2007). This statistics is calculated by locating the median of the entire

waveform, including both canopy and ground return energies, and computing the distance between this location and the ground return.

Given the use of relatively simple LiDAR metrics within this project, future work should involve the testing of more advanced LiDAR derived metrics in order to test if there is a relationship between them and forest structural and composition attributes. Some examples follow.

LiDAR derived point density metrics were not employed within this project due to differing point densities between areas of flight-line overlap. For example a canopy density metric is calculated using the proportion of returns whose height is 2 metres or greater divided by the total number LiDAR returns (Evans et al., 2009). Estimates of the ratio of penetration of LiDAR returns into the canopy can be calculated within a defined height bin, stratified by defined height ranges (Evans et al., 2009). This can be calculated as: $[n_i/n] \times 100$ (where: n_i = number of returns in height range i and n = total number of returns).

Canopy permeability is calculated as the proportion of laser pulses for which there are multiple returns, which is similar to the techniques used in Moffiet et al. (2005) to distinguish between poplar box and cypress pine tree species in South-East Queensland, Australia.

The foliage height diversity ratio metric is calculated as the proportion of returns in specified height strata. For example, Lesak et al. (2011) produced a number of metrics calculated from the proportions of LiDAR returns within 10 vertical height bins which related to bird species richness and forest structural types. One such example expanding upon these derived metrics is to calculate the index for return height diversity (H') by using the proportion of returns in each of the 10 equally spaced proportional density bins (p_i) (with volume defined by canopy height) by use of a modification to the SH diversity index, this can be expressed as:

$$H' = \sum_{i=1}^{10} p_i \log_{p_i} p_i \quad [10.1]$$

A relatively recent approach to the quantification of forest structure using LiDAR has been the ‘binning’ of the normalised 3D point cloud to reduce the data volume to a single measurement (Chasmer et al., 2004; Lee et al., 2004; Popescu and Zhao, 2008; Wang et al.,

2008). This method classifies height ranges within the canopy as a set of 3D volumetric pixels, or voxels. Such an approach has been applied to identify individual trees in 3D space (Lee et al., 2004), rather than plot or stand-level metrics. Attributes related to those individual trees such as crown dimensions and species (Vaughn et al., 2012) can be assessed through voxel cluster characteristics relating to their spatial arrangement and the LiDAR attributes of the points intersecting each voxel, such as the number of returns, in addition to amplitude and echo-width metrics should FW LiDAR data be available.

There are a number of additional metrics relating to topography and proximity to features which could potentially provide a benefit to the estimation of forest condition metrics. A number of examples are given in the research put forward by Martinuzzi et al. (2009) where 15 topographic metrics related to slope, aspect, spatial location, relief ratio, distance to water courses and a number of topographic indices were calculated and entered into the predictive modelling with 19 other LiDAR derived canopy metrics. The purpose of the approach was to detect and map deadwood and understorey. Likewise Simonson et al. (2013) incorporated spatial aggregation calculations using metrics calculated by FRAGSTATS for tree species richness and canopy height.

10.6.2.7 FW LiDAR pre-processing

It should be noted that some of the processing techniques applied to the FW LiDAR dataset were in development during the time of conducting this research project. One such feature which needs to be addressed is the consideration of base-line fitting in the individual waveforms recorded as part of this dataset. The base-line fitting was necessary for the reduction of spurious or false returns, or Gaussian peaks, from the waveforms. While the changes were often small, the base-line value for each waveform may have been different. Thus there is the small potential of over-fitting and under-fitting with the current pre-processing methodology.

Alternative methods exist for extracting 3D points from the FW LiDAR dataset other than using Gaussian functions. Mallet et al. (2009) and Chauve et al. (2007) argue that Gaussian functions are too general an approach which does not take into account the physical characteristics of the LiDAR waveforms and would not be suitable for complex forested landscapes. Mallet et al. (2009) review a number of other alternative stochastic functions based on marked point processes which showed promise.

As noted in section 10.3.3 (and figure 4.15) the FW LiDAR data exhibited anomalies further from nadir, typically in the overlap between flightlines. This problem was not explored during the course of this project, but needs to be in the future to understand the issues related to using FW data.

10.6.2.8 DR vs. FW LiDAR sampling disparity

A comparison of the relative merits of DR and FW LiDAR for the estimation of forest compositional and structural attributes was made as part of this thesis. In further analysis a comparison could be made if both acquisition parameters for DR and FW LiDAR were identical, either through new data acquisitions or modification of the existing datasets in order to create a somewhat fairer test. In addition, a quantification of the differences that the resolution of samples for the digitized return-waveforms for the FW LiDAR had on the ability to resolve metrics would be of great benefit for this project and the analysis of FW data in general.

10.6.2.9 Alternative statistical methods

The two regression approaches utilised within this thesis performed similarly in terms of producing predictive models and their accuracies, neither of which demonstrated a superior performance. OLS regression and AICc regression approaches are but two examples of statistically predicting attributes; there are many other regression approaches, such as least absolute shrinkage and selection operator, and alternatively methods such as neural networks, fuzzy logic models and regression trees. A linear relationship was also assumed between the field and remote sensing data. If for example additional sample points were collected and the relationships were determined to be non-linear a generalised additive model could be used to fit a spline to the relationship (Hastie, 2013).

There are also alternative methods to produce a subset or reduced number of input metrics for the predictive model, one such example is the use of Principle Component Analysis (PCA) transformation which is a technique for finding patterns in data of high dimensionality (Field, 2013). Other alternatives include shrinkage approaches, such as lasso regression (Friedman et al., 2013), to incorporate predictor subset selection.

A number of studies in the surrounding literature utilised alternative statistical methods to regression, one of the most prominent is that of the ‘random forest’ (RF) algorithm (Breiman, 2001). The RF algorithm is a nearest neighbour imputation, a form of non-parametric regression. The RF approach was not considered appropriate during the course of this thesis project due to an insufficient number of field plot measurements required for model training data, where approximately 80 plots would be required. Hudak et al. (2008) utilised such an approach with LiDAR data to estimate the forest structural attributes of stem density and basal area for 30x30m areas in Idaho, USA. Whereas Martinuzzi et al. (2009) utilised an RF approach with LiDAR data to estimate and map understorey presence/absence and the amount of standing deadwood again in Idaho, USA.

Another alternative approach, although dependent on accurate individual tree detection, would be the determination and application of allometric relationships to the extracted remote sensing metrics for each individual tree object. This approach is demonstrated in Lucas et al. (2008b) where individual trees were delineated and classified by species group. LiDAR derived heights and estimates of point densities were extracted. The application of species-specific allometric equations was then applied to estimate DBH and the above ground components of biomass.

All forest compositional and structural types were included in the statistical predictive modelling approach employed within this project. There are obvious differences in forest structure between coniferous and deciduous woodland types which may be better quantified by producing predictive models for each broad structural type rather than in general. A system could potentially be implemented in the future where broad species type would be identified, for example with hyperspectral data, and different sets of predictive equations applied.

10.6.2.10 Field data enhancement

As identified by Strunk et al. (2012) small numbers of training field plot data will have a deleterious effects upon model estimate precision and validity made from remote sensing data. A number of validation plots recorded as part of this research had poor correspondence with remote sensing model estimations because of dissimilarity to forest conditions enumerated in the initial training fieldwork. Thus, in order to reduce the likelihood of the

above, additional fieldwork would be advisable in future work for training data to improve the predictive statistical models overall.

The number of training field plots was small due to the time restraints and the requirements for detailed fieldwork. A pilot study could perhaps have been instituted to experiment with field metrics to be recorded and different plot sizes in order to make the process more streamlined. Future work should implement such an approach as the field data requirement was high.

10.6.2.11 Alternative condition assessment methods

Only a few examples of condition assessment techniques were tested in this study. There are a number of alternative assessment approaches considering individual trees and area-based calculations. One such approach was put forward by Van DemMeersschaut and Vandekerkhove (1998), where a score based index was calculated. This scoring index was calculated using a gradient of scores dependent on the similarity of field measurements to a number of desired forest compositional and structural conditions based upon Flemish forest inventory. A number of forest condition indices are based upon some manner of calculations involving field measurements, as summarised in McElhinny et al. (2005), Pommerening (2002, 2006), and Neumann and Starlinger (2001). One example is the stand variance index (STVI) (Staudhammer and LeMay, 2001) which is a combination of spatial diversity (arrangement) and tree attribute diversity, based on the covariance of tree height and DBH.

Given the findings in this project as regard to the conventional condition assessment indices, future work may produce a new condition assessment method which capitalises on the strengths of the current techniques while also minimising some of the problems encountered. The best method identified was the score-based technique which utilised 17 inputs to be assessed against target values. A number of improvements or modifications to this approach could be applied, two examples of which follow. Regenerating seedlings in an area may be below the set target for example, but this is not the same as an area having none at all, thus the target values could be changed in order incorporate a gradient of values, as in Van Dem Meersschaut and Vandekerkhove (1998). Tree size and tree size variation were important aspects of these condition indices, which in turn was linked to deadwood volume, within the context of this study. Potentially it should be possible to remove measures of deadwood in order to simplify the condition assessment.

10.6.2.12 Transferability

This particular study has demonstrated the potential for the estimation and mapping of forest compositional and structural parameters across a wide area within the New Forest. The application of this methodology to another woodland area in the future would provide a test ultimately of the transferability of this approach to other forest species and structural arrangements.

Many of the methods applied in pre-processing of the hyperspectral dataset, the classification approach, the calculation of the VIs and the extraction of LiDAR metrics could be applied to other areas given a similar dataset and assuming similar vegetation leaf-on and leaf-off conditions. Depending upon the similarity of the datasets and the environmental conditions, the statistical modelling of forest metrics may need to be re-implemented. The pre-processing techniques and VI calculation would be identical; however there are a number of issues relating to the application of the hyperspectral object-based classification method utilised in this project with regards to transferability.

The species classification approach employed an MNF transform upon the hyperspectral data to reduce the total number of image bands and reduce the amount of noise and extraneous data contained within each extent (Boardman and Kruse, 1994; Green et al., 1988). MNF imagery is an abstraction from the actual spectral measurements contained within the hyperspectral imagery, the digital number values produced do not relate to any scale. By its nature the resulting MNF imagery will be heavily influenced by the data acquisition scene, where different acquisitions will result in MNF imagery corresponding with different surface types, in other words two MNF datasets are not comparable.

Owing to the differences in MNF image data calculated between potential study sites the hierarchical classification with user defined membership rules developed for the New Forest study site will not be applicable to other areas. A new set of membership functions would have to be defined to account for these effects. In addition there are numerous potential methods of automatically detecting ITC objects from remote sensing data, although this is typically with LiDAR data (Kaartinen et al., 2012), which may better correspond with individual overstorey tree locations.

The transferability of the use of hyperspectral data for classification of trees would need to be considered or removed, as with the inclusion of LiDAR the significance of the classification map diminishes in terms of its importance in the predictive models constructed as part of this project.

Likewise the predictive regression models for all datasets may also need to be redefined where environmental conditions encountered were significantly different, or more specifically where these situations were not covered in the initial training field campaign and thus not represented in the statistical model.

10.7 Overview

This project sought to explore the possibilities of three airborne remote sensing techniques for the prediction of forest structural and composition attributes critical to the assessment of forest condition. Hyperspectral, DR LiDAR and FW LiDAR were tested independently and in combination (producing five datasets in total) for their ability to accurately estimate forest metrics using two regression techniques, OLS and AICc. These statistical models were derived by regressing the extracted remote sensing metrics against field measured values, (recorded in 2010). While there was no clear ‘best’ regression modelling approach, each of the datasets could predict many of the 39 field-level metrics when validated against field data recorded in 2012.

A total of 35 metrics out of a total of 39 could be estimated with NRMSE values below 0.4. The remaining four metrics could not be estimated either because of the regression approach failing to determine a relationship or the predictive model producing NRMSE values above 0.4. Many of the prediction accuracies were comparable with those published in the surrounding literature. Unfortunately not all the metrics could be found in the parallel remote sensing literature. Overall, out of the predictive models derived from each of the five remote sensing datasets, the regression models utilising FW LiDAR inputs produced the most accurate estimates of field metrics. The best models which used FW LiDAR inputs accounted for over half of the total 35 metric estimates.

Each of these predicted metrics could be mapped across the study site at the field plot-level, in this case 30x30m, utilising the validated models. Six conventional condition index assessment techniques were applied using the predicted metrics. The VE, Clark-Evans

aggregation and the diameter difference indices provided poor results due to issues related to underestimations of stem number from the automated ITC approaches or an insufficiently small range of index values. However, the HC, CSDI and score-based indices produced similar results, and were able to identify differences between forest types within the New Forest FC defined compartments. The score-based index combines elements from the HC and the CSDI indices in addition to many others, and avoids some of the challenges inherent in using ITC metrics directly. Overall this provides the most representative condition assessment method.

This thesis has provided a solution to the consistent retrieval of 35 metrics related to assessing forest condition. The methods developed in this research can be used to support both forestry and conservation assessments at scales from the field-plot (30x30m) to the landscape scale, in addition to demonstrating the operational use of FW LiDAR for forest research. The techniques described here have potential applications for satisfying ecological assessment criteria which require detailed knowledge of species composition, distributions and structure. The approach identified also has applications for predicting habitat types or suitability for both flora and faunal species. Many of the metrics extracted from the remote sensing data have immediate applications for precision forestry; for example tree number, species type, average DBH, variation in DBH size, and basal area can be predicted within a reasonable accuracy level (i.e. NRMSE < 0.4). If time series information were available the assessment of the impacts of various management activities though time could also be achieved.

There are a number of potential improvements which could be made in future work. For example, in terms of the addition of datasets which were unused, improvements which could be made to the ITC algorithms, the derivation of new remote sensing metrics, and the use of alternative statistical approaches.

Chapter 11 – Conclusions

11.1 Estimating forest metrics

This project has provided an evaluation of the integration of hyperspectral and LiDAR data sources in addition to the datasets in isolation, for estimating the numerous elements necessary for the assessment of forest condition for a study site in the New Forest, UK. In total there were three airborne remote sensing datasets, (i) hyperspectral, (ii) DR LiDAR and (iii) FW LiDAR, in addition to two combined datasets (iv) hyperspectral and DR LiDAR, and (v) hyperspectral and FW LiDAR. It is impossible to measure and enumerate every potential feature within the environment, thus a choice of condition indicators were selected through a review of the surrounding literature, which identified a number of forest structural and compositional metrics existing within both the overstorey and understorey. These metrics were recorded using fieldwork.

The remote sensing datasets were processed and area-based metrics extracted. The hyperspectral data included estimates of the number of overstorey tree species derived from ITC data and numerous area-based vegetation indices calculated from the spectral data, which related to the volume of green vegetation, light-use efficiency, stress-related pigments and water content within the forest canopy. The DR LiDAR data had numerous area-based metrics generated relating to the distribution of heights and return intensity values, in addition to summarised ITC-metrics. The FW LiDAR data required additional processing steps. This included the fitting of Gaussian peaks to each of the returned waveforms and extracting 3D point information. Once completed, FW LiDAR had numerous metrics extracted relating to the distribution of heights, echo-amplitude and echo-width values, in addition to summary ITC-metrics.

Species classification of delineated ITC objects was performed using an object-based hierarchical classification method with used defined membership class functions. This approach produced mapping accuracy of 62% for 26 classes. This accuracy level compares well with other similar approaches (Heinzel and Koch, 2013; Lucas et al., 2008b). A number of issues with over- and under-segmentation were observed because of the ITC delineation method which will influence mapping accuracy level.

Statistical regression models were applied to establish relationships between field measured structural and compositional features for each of the three remote sensing datasets and two combinations (hyperspectral combined with DR LiDAR, and hyperspectral combined with FW LiDAR). The regression model prediction results were then assessed by comparing them with field recorded values. This study attempted to estimate a total of 39 field plot-level metrics, whereas many studies identified in the surrounding literature focus on extracting relatively few.

The predictive models derived from hyperspectral data could predict 19 out of a total of 39 field plot-level metrics with a reasonable accuracy ($\text{NRMSE} < 0.4$), whereas only three estimates had NRMSE below a value of 0.3. Unsurprisingly many of the metrics which were related to structural metrics, such as basal area, or volume of deadwood were estimated poorly. The majority of these models utilised VI input metrics.

ITC analysis was conducted using LiDAR data and was utilised both in the tree species classification and for estimation of stem density and tree crown attributes. Object-based tree species classification was conducted using the ITC objects and a dimensionally reduced hyperspectral dataset. It was found that the ITC method deployed within this thesis typically underestimated the total number of tree stems and overstorey tree stems within an area, especially so in field plot areas with a high stem density. However estimates of crown dimensions corresponded well with field data estimates, for example crown diameter varied from 1.8-4.0m RMSE.

A number of metrics could be estimated directly from DR LiDAR, such as average canopy height, the number of canopy layers and canopy openness, in addition to those derived by regression models. The model estimations derived from DR LiDAR produced 30 metric estimates with NRMSE values below 0.4, and 16 below 0.3. Values with a higher RMSE were typically related to species composition, in particular that of the understorey. Many of the regression models produced utilised inputs from ITC data, in particular the ITC-metrics relating to the average and total crown dimensions per plot, in addition to other secondary metrics from point cloud statistics. This model behaviour is consistent with Hyypäet al. (2001), Person et al. (2002), Popescu et al. (2004) and Maltamo et al. (2004).

The results of the model estimates for forest structure related metrics corresponded well with those recorded in the surrounding literature. These components were the stem density, average DBH, total basal area, height to the living crown, the average crown horizontal area and the total crown area within the plot. Many of these metrics are required for operational forest inventory, and produced similar RMSE values to these approaches (Hudak et al., 2009; Hyypä et al., 2008; Kaartinen et al., 2012; Lim et al., 2003a; Naesset et al., 2004; Richardson and Moskal, 2011). Estimates of the variability of these metrics were also calculated. Estimation of the basic metrics of overstorey composition produced relatively small RMSE values using DR LiDAR only. These predictive models employed a combination of DR LiDAR height and intensity value distributions. Error values for estimates relating to Shannon and Simpson indices were high relative to the index ranges however.

Estimates of understorey structure and composition derived from DR LiDAR generally produced models with relatively low RMSE values. Standing and fallen deadwood volume and the decay level had NRMSE values below 0.37. Estimates of regenerating species number and species generally had NRMSE values below 0.34, apart from the number of sapling species and the Shannon index for native seedlings species which had large RMSE values or could not be modelled. The estimates for the number of vascular species and the percentage cover of ground vegetation had NRMSE values below 0.46. All of these metrics utilised a combination of vertical and intensity metrics.

FW LiDAR was acquired at the same time as the DR, utilising the same scanning system. The initial sampling parameters for FW LiDAR were approximately 40% of the total pulses acquired for DR LiDAR. Through Gaussian fitting applied to the returned laser waveform information, a larger number of additional 3D points could be derived. Through this procedure the FW LiDAR generated an additional 143% 3D points from the multiple returns in relation to the initial pulses. DR LiDAR only generated an additional 30% of points from the initial pulses. The Gaussian decomposition method applied to the FW LiDAR data produced similar numbers of total points to that of the DR LiDAR data overall. Similar distributions of points were also present within the forest canopy, however FW datasets generally produced a higher number of returns within the mid-canopy.

As with DR LiDAR, relevant FW metrics for assessing woodland condition could be estimated directly, such as average canopy height, canopy openness and the number of

canopy layers. The estimates of average canopy height derived from DR and FW LiDAR were very similar, with RMSE values of 3.5 and 3.7m respectively. The estimates for canopy openness for DR gave an RMSE value of 13.29%, whereas the FW value was 8.28%. The estimates for the number of canopy layers for DR gave an RMSE of 1.2 layers, whereas the FW value produced a value of 1.1 layers.

The regression model estimates of the condition indicator field plot-level metrics derived from FW LiDAR produced a total of 26 metric predictions with NRMSE below 0.4 and 18 below 0.3. Again, many FW LiDAR models utilised predictive inputs from ITC crown dimension metrics. The additional FW-based metrics relating to echo-width also saw a wide usage for many structural and compositional components, an observation also made in Reitberger et al. (2008).

In comparison many of the RMSE values for the field-level metrics estimated were similar between DR and FW LiDAR overall. The FW model estimates for forest structural components, such as average DBH and height to the living crown, compared favourably to a number of studies available in the surrounding literature for DR LiDAR. Estimates of the variability of these metrics were also calculated. The estimates of overstorey composition produced similar RMSE values for both DR and FW LiDAR. FW produced slightly lower error for estimates of the number of native trees and the number of tree species. The number of tree species utilised both amplitude and echo-width metrics from a vertical level within the canopy, which supports Reitberger et al. (2008) who stated that combining geometric, amplitude and width metrics improves accuracies of species classification.

The majority of regression model estimates of understorey structure and composition derived from FW LiDAR had low RMSE values, with a few exceptions. The standing and fallen deadwood volume estimates produced RMSE values of 0.17m^3 and 0.27m^3 , respectively. These values were both slightly smaller than the equivalent estimates made using DR LiDAR. The estimate of decay level for standing deadwood was better for FW (RMSE 0.19) than DR LiDAR (RMSE 0.24), whereas the decay level estimate for downed deadwood was slightly better for DR (RMSE 0.17) in comparison to FW LiDAR (RMSE 0.20). FW LiDAR derived estimates of understorey regenerating tree density and species types were very similar to DR, including an estimate of the number of sapling species. FW LiDAR could not produce a model estimation for the Shannon index for native seedling species.

The benefits of fusion of (i) hyperspectral and DR LiDAR, or (ii) hyperspectral and FW LiDAR metrics for modelling field plot-level metrics was assessed in relation to the best models, i.e. the estimations with the smallest overall RMSE, produced from either hyperspectral, DR LiDAR or FW LiDAR in isolation from one another. A total of nine models from both composite datasets were found to better estimate field-plot values. Apart from the estimates of total crown horizontal area, each of the others was related to the estimation of overstorey and understorey compositional attributes.

Over the course of this project a fusion of the two dataset types was applied multiple times, for example the object-based classification of tree species and the area-based metric estimation. In order to produce the most complete list and best estimated of the total 39 field plot-level metrics it was necessary to apply the regression analysis to all five datasets. Two relatively minor metrics could not be modelled at all; these were the percentage of native saplings and the number of bryophyte species. However two model estimations produced very poor estimates, these were the percentage of very big trees (DBH >80cm) and the Shannon index (by basal area). For the 35 remaining metrics, some metrics could only be extracted from either hyperspectral or LiDAR data, or a combination. Some datasets alone could not produce estimates at all. The wealth of information the cumulative datasets provides accounts for the shortcomings of the individual datasets. In particular, FW LiDAR provided the most input to predicating 20 (15 LiDAR and 5 composite) (57%) of the field-level metrics, in comparison to 13 for DR (9 LiDAR and 4 composite) (37%), and 2 for hyperspectral models (6%). Nonetheless overall RMSE values were similar between both DR and FW LiDAR, and in some cases hyperspectral estimates as well. The metrics that the FW LiDAR best predicts account for over half the total 35 models. This again confirms the benefits of the additional metrics the FW datasets provides, in particular the echo-width metrics.

FW LiDAR data provides a benefit to predicting a significant proportion of the forest condition field plot-level metrics addressed over the course of this research, and which has proven to be at least as good as DR LiDAR based estimates, especially when considering the initial disparity between the FW and DR LiDAR datasets in terms of spatial sampling. The FW models utilise metrics related to echo-width distribution extensively, an observation also made by Reitberger et al. (2008), which may indicate the explanatory power the additional metrics derived from waveforms may provide. Additionally, the predicative power of ITC

crown average, standard deviation and total horizontal area was very important for estimating many structural metrics, which is supported by other studies (Hyyppä et al., 2001; Person et al., 2002; Popescu et al., 2004; Maltamo et al., 2004).

11.2 Methods of assessing condition

A total of six condition indices were identified in the surrounding research literature, these were: (i) the Clark-Evans aggregation index (Clark and Evans, 1954); (ii) the tree diameter differentiation index (Földner, 1995, cited in Vorčák et al., 2006); (iii) the vertical evenness index (VE) (Neumann and Starlinger, 2001); (iv) the complexity index (HC) (Holdridge, 1967); (v) the complex stand diversity index (CSDI) (Jaehne and Dohrenbusch, 1997, cited in Neumann and Starlinger, 2001; Vorčák et al., 2006); and (vi) the scoring method as defined by Cantarello and Newton (2008). Each of these indices was calculated using remote sensing derived estimates of their required inputs for areas corresponding with the validation field plots and for the whole study site for mapping purposes.

The VE and aggregation indices were computed from ITC delineated data and were validated against field data calculated indices. Owing to the stem number underestimation encountered with the ITC delineation both of these indices provided poor results. In the future it may be possible to derive proxies for these indices using the LiDAR point cloud data itself, such as in Lesak et al. (2011). The remaining four indices utilise area-based metrics derived statistically, apart from the CSDI which utilises both statistically derived and ITC-metrics in its construction.

The diameter differentiation index was calculated from two statistically derived area-based metrics; stem density and the sum of the DBH differences per plot. The predicted index values performed well with the field data, however in the mapping product there was little difference between values across the plot and so this proved to be a poor estimate of condition.

The HC index was calculated using four area-based inputs; the stem density, number of tree species, average canopy height and total basal area. The predicted index value corresponded reasonably with that measured in the field (i.e. within 20%), however field plots with high stem density suffered underestimations. The HC index is more indicative of tree size

parameters; the mapping product also tends towards an almost binary output depicting coniferous or deciduous woodland types.

The CSDI was the most computationally complex method for assessing site condition. The index was computed from the additive sum of four component indices each assessing a different element of forest structure or composition; these were (i) species diversity; (ii) DBH diversity; (iii) stem spacing diversity; and (iv) tree crown diversity. A combination of ITC and area-based metrics were used as input. While there were limitations with some of the ITC-metrics, the index value corresponded well with the field data calculated index. The CSDI mapping product conformed well to known structural divisions within the study site extent. The CSDI relates more to structural diversity which can transcend either the broad types of deciduous or coniferous woodland, however problems occurred in grid-cells containing large canopy gaps which are of concern.

The final condition assessment method utilised an approach of additive scoring based on whether 17 field metrics reached a desired target value. An ideal site would receive a maximum score of 17. Overall the remote sensing metrics combined within this index corresponded well to those derived from field data. Again the score-based index mapping product conformed well to known structural divisions within the study site extent. The score-based index combines both size and diversity metrics at a relatively basic level in addition to many other factors. This method was considered the most informative and contained fewer uncertainties because of not including ITC-metrics directly.

The calculation of condition assessment indices involved error from each of the input estimated metrics. While this is present, the additional uncertainties proved to be small. The results presented indicate that it is possible to produce estimates of forest condition from remote sensing products utilising conventional methods of assessment.

The six condition indices were assessed against each other by initially categorising the index scales to make them consistent with one another and re-mapping the resultant values per 30x30m grid-cell, as in Chapter 9.3. The categorisation highlighted an issue with the diameter differentiation index, where the range of predicted values was too small to make meaningful assessments. The VE and Clark-Evans aggregation indices were also poor; this was due to problems with ITC underestimation of trees within a plot, which was more severe in

deciduous areas with a dense and closed canopy. The remaining three indices however (HC, CSDI and score-based) showed distinct spatial patterns which corresponded well to known structural types and FC compartment boundaries within the New Forest study site. In order to compare the HC, CSDI and score-based indices a mean and standard deviation were calculated for each of the mapped grid-cells.

When comparing the HC, CSDI and score-based indices mean and categorised maps, a great deal of similarity between the three techniques was observed. Each of the indices gave high values to deciduous areas, which were dominated primarily by oak and beech trees, and were commonly over 70-100 years old. Only very few coniferous areas exhibited consistently high index values. The areas with high index values were defined by the presence of large trees (i.e. DBH values) with variable sizes, in addition to generally higher numbers of tree species. Often these features were accompanied by larger volumes of both standing and fallen deadwood. Areas such as these may also contain understorey components; however this was variable in areas of high index values for the HC index and CSDI. Understorey components across much of the site were low however. The majority of coniferous dominated compartments exhibited low values from all indices.

When interrogating the standard deviation of the three index values across the study site map it is clear the HC and CSDI reflect different elements of the composition and structure of woodland, especially when applied to deciduous woodlands, such as Frame Wood in particular. The HC index is more sensitive to large tree sizes and tends towards an almost binary output for depicting coniferous (low) or deciduous (high) areas. The CSDI on the other hand related to species and structural diversity, which produced generally higher index values in deciduous areas which can also be more variable, including low to high values. The CSDI has an issue when estimating values in grid-cells which contain large canopy gaps. The score-based index combined both tree size and diversity metrics at a relatively basic level in addition to many other factors, such as deadwood and understorey. The score-based index also did not directly use ITC-metrics and thus should be less influenced by its associated error.

A provisional assessment of the woodland condition within the New Forest was carried out using the score-based index method, as defined in Cantarello and Newton (2008). The index value was utilised as a guide to identify spatial patterns within the mapping data with areas of

good or poor condition status. The maximum index value recorded was 14, so none of the sites can be considered to be ‘perfect’. The highest values occurred in semi-ancient deciduous woodlands while many coniferous compartments exhibited low values, because of deficiencies in species number, deadwood amount and understorey.

In terms of forest management it is clear that current management operations have a high impact on the structural and composition components within the New Forest site, causing similar attributes within each compartment caused by distinct planting dates, understorey regeneration is low or non-existent because of little ungulate control, and deadwood resources are removed in plantation areas by FC workers. Each of these elements has an impact on overall condition, thus the above analysis highlights some of the shortcomings of current management approaches.

11.3 Key findings

Throughout the course of this project the analysis of the three airborne remote sensing datasets has provided a means of extracting a wealth of information about forested environments. The hyperspectral, DR LiDAR and FW LiDAR were analysed in isolation of one another and could produce estimates of various forest compositional and structural components at the field plot-level.

For the analysis of hyperspectral datasets an object-based species classification was performed to estimate species number, species variety and relative abundance. The accuracy levels were comparable with those published in other studies, but did contain a number of errors due to the ITC delineation procedure implemented. In terms of the hyperspectral data inputs into the statistical models for estimating the area-based plot-level forest metrics the importance of the object-based species classification was relatively low in comparison to the input metrics relating to vegetation indices (VI).

There were two LiDAR datasets analysed in the course of this study, DR and FW. The initial sampling of the two datasets was higher for DR LiDAR, however the further processing of the FW LiDAR through Gaussian fitting provided as many total returns as DR LiDAR. Overall the DR and FW LiDAR statistical model predictions produced similar accuracies, which were in line with those published in the surrounding literature. In comparison the statistical model inputs for FW LiDAR often utilised fewer or more appropriate inputs for the

prediction of understorey field plot-level metrics. Many of the FW LiDAR predictive models also utilised metrics relating to the echo-width of the returns. While there are still a number of uncertainties about the properties of these new FW LiDAR metrics (Mills, 2010; Wagner et al., 2006), their influence on the results offers promise for future work. Nonetheless this project demonstrates the use of FW LiDAR which offers comparable if not better results to those of DR LiDAR.

The best models (i.e. with the lowest RMSE) estimated overall for all of the required forest compositional and structural metrics came from a variety of the three datasets individually and in combination. A total of 35 of 39 field plot-level (30x30m) forest metrics could be estimated, many of which when validated produced NRMSE values below 0.4, and 23 models produce NRMSE values below 0.3. Over half of these models involved the use of FW LiDAR data on its own or FW combined with hyperspectral data. These models could be applied across the whole study site producing a map for each of the 35 field metrics.

The focus of this project was to utilise airborne remote sensing to assess forest condition. To this end, six conventional condition assessment indices were tested with the derived remote sensing metrics. Three of which proved to be poor, the remaining three however corresponded well with field derived index values and when mapped across the study site extent corresponded well with known structural and compartment boundaries. Deciduous areas received higher index values, and coniferous areas typically received low values. From comparing these index results, tree size and differentiation are inherent with high index values, and where these higher index values are present a higher proportion of both standing and downed deadwood are found. Understorey regenerating tree and vascular plant species population and number are typically not linked to tree size or differentiation within a plot, indicating that in the context of this research the score-based method is the most representative method of assessment. The score-based forest condition assessment method was applied to the New Forest study site and confirmed known issues with current management objectives, such as there being little or no regeneration of tree species throughout the site, in addition to coniferous compartments exhibiting structural and compositional homogeneity, and very low levels of deadwood.

The approach demonstrated in this project shows that conventional methods of assessing forest condition can be applied with remote sensing derived inputs for woodland assessment

purposes. This procedure was applied to a large spatial extent and provided a detailed assessment, which at this scale would be impractical with field data alone. The approach is quantifiable, with known levels of accuracy, and is less subjective than some current field-based methods of assessing condition, such as the researcher answering a questionnaire while following a structured walk.

11.4 Key conclusions

The following are the key conclusions of this research:

1. The most relevant field plot measurements are tree species, DBH, crown horizontal area, average tree height, the number of saplings, number of seedlings and vascular plant diversity. All of the other field metrics can be estimated from these metrics;
2. Individual tree crown (ITC) delineation can only be considered as semi-operational due to poor segmentation results;
3. Object-based classification of hyperspectral data can provide sufficient accuracy for the mapping of species type, number and proportions of native species;
4. Metrics from the hyperspectral species classification are less valuable than vegetation index (VI) data for estimation of forest metrics;
5. Hyperspectral remote sensing data are optimal for tree species composition mapping, but poor for estimating structure;
6. LiDAR data are critical for estimating forest structural metrics;
7. DR and FW LiDAR models estimate forest structure with similar levels of accuracy, although generally FW LiDAR models are slightly higher;
8. Applying Gaussian decomposition to FW LiDAR data generates ca. three times the number of points per pulse than for DR, which mitigates the lower pulse rate of the FW system;
9. Echo-width metrics from the FW LiDAR are important additional variables for forest metric predictions unavailable in DR LiDAR models;
10. A combination of hyperspectral and LiDAR data is best for estimating forest understorey component metrics, such as the number of seedlings;
11. LiDAR derived ITC-metrics relating to crown horizontal area are important inputs in predictive models of forest structure;
12. Other ITC-metrics however, such as estimates of stem number and spacing, are currently poor and should not be used for modelling;

13. Of six condition assessment approaches tested, only three function well (complexity index, complex stand diversity index, score-based index);
14. The three successful condition assessment methods are driven primarily by tree size and tree size variation;
15. The presence of standing and downed deadwood is positively linked to increases in tree size and tree size variation, whereas the presence of understorey is independent from this;
16. The best technique for assessing woodland condition is the score-based method, this combines seventeen inputs which relate to tree species composition, tree size and variability, deadwood, and understory components;
17. All of the above woodland characteristics can be derived from the appropriate combination of airborne hyperspectral and LiDAR datasets with an accuracy of at least $\text{NRMSE} < 0.4$;
18. Therefore it has been shown possible to map forest condition over landscape-scale areas by airborne remote sensing.

11.5 Final thoughts

The approach produced for this project demonstrated that it is possible to estimate a range of forest metrics from three types of airborne remote sensing datasets for both structural and compositional metrics across a landscape at a level unachievable through plot-based survey alone. While estimation of both individual tree and area-based metrics is an involved process and there is clear potential for improvement, the approach demonstrates advantages with applying remote sensing data in such a manner to generate continuous estimates of condition metrics and indices across a wide area with a relatively low error through the combination of both hyperspectral, DR and FW LiDAR data. Many of the forest metrics estimated are either directly comparable with that presented in the field, or have not been previously attempted, many of which were estimated with relatively high accuracy.

The approaches developed for the characterisation of forest composition and structure of the selected New Forest site have direct applications for forest condition assessment, habitat suitability modelling and forestry in other regions of the UK and overseas. The transferability of the use of hyperspectral data for classification of tree species would need to be considered or removed, as the significance of the classification map diminishes in terms of its importance in the predictive models constructed with the inclusion of LiDAR datasets. The overall approach identified, however, could be potentially applied in many environmental contexts, as similar approaches exist in the literature for the estimation of forest condition from airborne remote sensing, especially so for DR LiDAR, for applications across the planet.

In future work further refinements to the approach could be implemented and tested. Refinements could be made to the ITC delineation method, which, for example, would improve/augment estimates of stem density, stem spacing, and object-based species classifications. Similarly the detection and mapping of suppressed trees would be beneficial. Both hyperspectral and LiDAR metrics could be incorporated to a higher degree for tree species classifications. There are also many more potential LiDAR point cloud statistics which may be of use to future analysis, for example point densities in particular vertical strata. Metrics generated from FW LiDAR datasets offer a great deal of explanatory power, future work may be able to produce more metrics such as the backscatter cross-section or coefficient. In addition supplementary data such as the area's proximity to watercourses or footpaths could be incorporated, for example as a proxy for levels of disturbance.

There are also a number of alternative approaches for the estimation of field plot-level metrics, for example the random forest algorithm. In accordance with sampling design more fieldwork samples would potentially improve the precision and validity of estimates. There are also a number of alternative conventional forest condition indices in existence which could be explored, adapted and applied, dependent upon the future projects objectives (e.g. McElhinny et al., 2005; Neumann and Starlinger, 2001; Pommerening, 2002; 2006).

This project demonstrates that a large number of metrics from a combination of remote sensing datasets can be extracted for the assessment of forest condition. A wealth of information is available for the composition and structure of the forest, including deadwood, understorey composition and regenerating tree species. The approach demonstrated here

could potentially be applied to estimate other forest attributes, for example estimating total above-ground biomass. Each of these forest attributes can be mapped across large areas. These remote sensing derived metrics can be input into conventional techniques for assessing forest condition, the output of which can be mapped. While the requirement for fieldwork remains, the approach as demonstrated in this project provides a far more detailed and/or comprehensive indication of condition than is possible through field work assessment alone in terms of spatial extent, and on a scale appropriate to observe spatial patterns of features within the stand-level.

Cited references:

- ARSF. 2011a. Full waveform notes. [Online] [Available from: <<http://arsf-dan.nerc.ac.uk/trac/wiki/Sensors/LeicaLIDAR/FWMarksNotes>>] [Date accessed: 15/03/2011].
- ARSF. 2011b. ARSF Data Processing. [Online] [Available from: <<http://arsf.nerc.ac.uk/data/>>] [Data accessed: 02/02/2013]
- Acker, S. A., Sabin, T. E., Ganio, L. M., and McKee, W. A., 1998. Development of old-growth structure and timber volume growth trends in maturing douglas-fir stands. *Forest Ecology and Management*, 104 (1), 265-280.
- Adams, T., Beets, P., and Parrish, C., 2012. Extracting more data from lidar in forested areas by analyzing waveform shape. *Remote Sensing*, 4 (3), 682-702.
- Alexander, C., Tansey, K., Kaduk, J., Holland, D., and Tate, N. J., 2010. Backscatter coefficient as an attribute for the classification of full-waveform airborne laser scanning data in urban areas. *Isprs Journal of Photogrammetry and Remote Sensing*, 64 (5), 423-432.
- Andersen, H. E., Mcgaughey, R. J., and Reutebuch, S. E., 2005. Estimating forest canopy fuel parameters using lidar data. *Remote Sensing of Environment*, 94 (4), 441-449.
- Andersen, H. E., Reutebuch, S. E., and Mcgaughey, J. H., 2006. (chapter 3) *active remote sensing* Springer.
- Anderson, J. E., Plourde, L. C., Martin, M. E., Braswell, B. H., Smith, M. L., Dubayah, R. O., Hofton, M. A., and Blair, J. B., 2008. Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest. *Remote Sensing of Environment*, 112 (4), 1856-1870.
- Aplin, P., 2005. Remote sensing: Ecology. *Progress in Physical Geography*, 29 (1), 104-113.
- Armston, J. Rsc lidar tools. Available from: <http://code.google.com/p/rsclastools/> [Accessed: 23/05/2012].
- Asner, G. P., 1998. Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment*, 64 (3), 234-253.
- Barton, K. 2012. Multi-model inference. [Online] [Available from: <<http://cran.r-project.org/web/packages/MuMIn/MuMIn.pdf>>] [Accessed: 12/04/2012].
- Bater, C. W., and Coops, N. C., 2009. Evaluating error associated with lidar-derived dem interpolation. *Computers & Geosciences*, 35 (2), 289-300.
- Baulies, X., and Pons, X., 1995. Approach to forestry inventory and mapping by means of multispectral airborne data. *International Journal of Remote Sensing*, 16 (1), 61-80.
- Blackburn, G. A. and Milton, E. J., 1996. Filling the gaps: Remote sensing meets woodland ecology. *Global Ecology and Biogeography Letters*, 5(4-5): 175-191
- Boardman, J. W., And Kruse, F. A., . 1994, . *Automated spectral analysis: A geological example using aviris data, north grapevine mountains, nevada*:. Paper presented at the Proceedings, ERIM Tenth Thematic Conference on Geologic Remote Sensing, , Environmental Research Institute of Michigan, Ann Arbor, MI.

- Bork, E. W., and Su, J. G., 2007. Integrating lidar data and multispectral imagery for enhanced classification of rangeland vegetation: A meta analysis. *Remote Sensing of Environment*, 111 (1), 11-24.
- Borre, J. V., Paelinckx, D., Mucher, C. A., Kooistra, L., Haest, B., Blust, G. D., and Schmidt, A. M., 2011. Integrating remote sensing in natura 2000 habitat monitoring: Prospects on the way forward. *Journal for Nature Conservation*, 19 (2), 116-125.
- Boudreau, J., Nelson, R. F., Margolis, H. A., Beaudoin, A., Guindon, L., and Kimes, D. S., 2008. Regional aboveground forest biomass using airborne and spaceborne lidar in quebec. *Remote Sensing of Environment*, 112 (10), 3876-3890.
- Brandtberg, T., Warner, T. A., Landenberger, R. E., and McGraw, J. B., 2003. Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in north america. *Remote Sensing of Environment*, 85 (3), 290-303.
- Breiman, L., 2001. Random forests. *Machine Learning*, 45 (1), 5-32.
- Broughton, R., Hinsley, S., Bellamy, P., Hill, R., and Rothery, P., 2006. Marsh tit poecile palustris territories in a british broad-leaved wood. *IBIS*, 148 (4), 744-752.
- Bunting, P., Armston, J., Clewley, D., and Lucas, R., 2011a. The sorted pulse software library (spdlib): Open source tools for processing lidar data. *Proceedings of SilviLaser 2011, 11th International Conference on LiDAR Applications for Assessing Forest Ecosystems*, University of Tasmania, Australia, 16-20 October, pp. 1-11
- Bunting, P., Armston, J., Lucas, R., and Clewley, D., 2011b. Sorted pulse data (spd) format: A new file structure for storing and processing lidar data. *In: SilviLaser, Oct. 17-19*, Hobart, Australia.
- Bunting, P., and Lucas, R., 2006. The delineation of tree crowns in australian mixed species forests using hyperspectral compact airborne spectrographic imager (casi) data. *Remote Sensing of Environment*, 101 (2), 230-248.
- Burnham, K. P., and Anderson, D. R., 2002. *Model selection and multimodel inference: A practical information-theoretic approach (2nd edition)*. New York: Springer-Verlag.
- Burrough, P. A., and McDonnell, R. A., 1998. *Principles of geographical information systems*. New York: Oxford University Press.
- Campbell, J., and Wynne, R. H., 2011. *Introduction to remote sensing (5th edition)*. London & New York: Taylor & Francis.
- Cantarello, E., and Newton, A. C., 2008. Identifying cost-effective indicators to assess the conservation status of forested habitats in natura 2000 sites. *Forest Ecology and Management*, 256 (4), 815-826.
- Cescatti, A., 2007. Indirect estimates of canopy gap fraction based on the linear conversion of hemispherical photographs - methodology and comparison with standard thresholding techniques. *Agricultural and Forest Meteorology*, 143 (1-2), 1-12.
- Chang, C., 2007. *Hyperspectral data exploitation: Theory and applications*. Hoboken, New Jersey: John Wiley & Sons.
- Chasmer, L., Hopkinson, C., and Treitz, P., 2004. Assessing the three-dimensional frequency distribution of airborne and ground-based lidar data for red pine and mixed deciduous forest plots. *In: ISPRS International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*.

- Chatters, C., and Read, M., 2006. *New forest national park*. Tiverton, Devon: Halsgrove.
- Chauve, A., Mallet, C., Bretar, F., Durrieu, S., Deseillgny, M., and Puech, W., 2007. Processing full-waveform lidar data: Modelling raw signals. *Internation Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*.
- Chen, J. M., 1996. Canopy architecture and remote sensing of the fraction of photosynthetically active radiation absorbed by boreal conifer forests. *IEEE Transactions on Geoscience and Remote Sensing*, 34 (6), 1353-1368.
- Chen, Q., Baldocchi, D., Gong, P., and Kelly, M., 2006. Isolating individual trees in a savanna woodland using small footprint lidar data. *Photogrammetric Engineering and Remote Sensing*, 72 (8), 923-932.
- Chen, Q., Gong, P., Baldocchi, D., and Tian, Y. Q., 2007a. Estimating basal area and stem volume for individual trees from lidar data. *Photogrammetric Engineering and Remote Sensing*, 73 (12), 1355-1365.
- Chen, Q., Gong, P., Baldocchi, D., and Xie, G., 2007b. Filtering airborne laser scanning data with morphological methods. *Photogrammetric Engineering and Remote Sensing*, 73 (2), 175-185.
- Chen, X., Ender, P., Mitchell, M., and Wells, C. Regression with spss. Available from: <http://www.ats.ucla.edu/stat/spss/webbooks/reg/default.htm> [Accessed: 12/08/2012].
- Christensen, M., Heilmann, C. J., Walley, R., and Adamcik, S., 2004. Wood-inhibiting fungi as indicators of nature value in european beech forests. *EFI Proceedings*, 51, 229-238.
- Clark, F. J., and Evans, F. C., 1954. Distance to nearest neighbour as a measure of spatial relationship in populations. *Ecology*, 35, 445-453.
- Clark, M. L., Roberts, D. A., and Clark, D. B., 2005. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment*, 96 (3-4), 375-398.
- Clawges, R., Vierling, K., Vierling, L., and Rowell, E., 2008. The use of airborne lidar to assess avian species diversity, density, and occurrence in a pine/aspen forest. *Remote Sensing of Environment*, 112 (5), 2064-2073.
- European Commission, 2010, EU biodiversity indicators – sebi 2010 [Online] [Available from: http://ec.europa.eu/environment/nature/knowledge/eu2010_indicators/index_en.htm] [Accessed: 05/07/2012].
- Coops, N. C., Wulder, M. A., Culvenor, D. S., and St-Onge, B., 2004. Comparison of forest attributes extracted from fine spatial resolution multispectral and lidar data. *Canadian Journal of Remote Sensing*, 30 (6), 855-866.
- Culvenor, D. S., 2002. Tida: An algorithm for the delineation of tree crowns in high spatial resolution remotely sensed imagery. *Computers & Geosciences*, 28 (1), 33-44.
- Curran, P. J., Windham, W. R., and Gholz, H. L., 1995. Exploring the relationship between reflectance red edge and chlorophyll concentration in slash pine leaves. *Tree Physiology* 15, 203-206.
- Dalponte, M., Bruzzone, L., and Gianelle, D., 2008. Fusion of hyperspectral and lidar remote sensing data for classification of complex forest areas. *IEEE Transactions on Geoscience and Remote Sensing*, 46 (5), 1416-1427.

- Dash, J., and Curran, P. J., 2004. The meris terrestrial chlorophyll index. *International Journal of Remote Sensing*, 25 (23), 5403-5413.
- Datt, B., 1999. A New Reflectance Index for Remote Sensing of Chlorophyll Content in Higher Plants: Tests Using Eucalyptus Leaves. *Journal of Plant Physiology* 154, 30-36.
- Daughtry, C. S. T., Hunt Jr., E. R., and McMurtrey Iii., J. E., 2004. Assessing crop residue cover using shortwave infrared reflectance. *Remote Sensing of Environment* 90, 126-134.
- Devictor, V., Julliard, R., and Jiguet, F., 2008. Distribution of specialist and generalist species along spatial gradients of habitat disturbance and fragmentation. *Oikos*, 117, 507-514.
- Drake, J. B., Knox, R. G., Dubayah, R. O., Clark, D. B., Condit, R., Blair, J. B., and Hofton, M., 2003. Above-ground biomass estimation in closed canopy neotropical forests using lidar remote sensing: Factors affecting the generality of relationships. *Global Ecology and Biogeography*, 12 (2), 147-159.
- Envi-Online-Help. 2005. Envi reference guide: Envi routines. [Online] [Available from: <http://geol.hu/data/online_help>] [Accessed: 15/07/2012].
- Esri. 2013 Arcgis desktop 10 help. [Online] [Available from: <<http://resources.arcgis.com/content/arcgisdesktop/10.0/about>>] [Accessed: 07/01/2013].
- Estornell, J., Ruiz, L. A., Velazquez-Marti, B., and Fernandez-Sarria, A., 2011. Estimation of shrub biomass by airborne lidar data in small forest stands. *Forest Ecology and Management*, 262 (9), 1697-1703.
- Evans, J. S., Hudak, A. T., Faux, R., and Smith, A. M. S., 2009. Discrete return lidar in natural resources: Recommendations for project planning, data processing, and deliverables. *Remote Sensing*, 1 (4), 776-794.
- Falkowski, M. J., Evans, J. S., Martinuzzi, S., Gessler, P. E., and Hudak, A. T., 2009a. Characterizing forest succession with lidar data: An evaluation for the inland northwest, USA. *Remote Sensing of Environment*, 113 (5), 946-956.
- Falkowski, M. J., Wulder, M. A., White, J. C., and Gillis, M. D., 2009b. Supporting large-area, sample-based forest inventories with very high spatial resolution satellite imagery. *Progress in Physical Geography*, 33 (3), 403-423.
- Faraway, J. 2011. R package 'faraway' version 1.0.5 (2011-04-05 ed.) [Online] [Available from: <<http://cran.r-project.org/web/packages/faraway/>>] [Date accessed: 07/07/2012]
- Field, A. P., 2013. *Discovering statistics using ibm spss statistics: And sex and drugs and rock 'n' roll (fourth edition)*. London: Sage publications.
- Florence, R. G., 2004. *Ecology and silviculture of eucalypt forests*. CSIRO Publishing.
- Forestry-Commission. 2012. About the new forest. [Online] [Available from: <<http://www.forestry.gov.uk/website/forestry.nsf/byunique/infd-69vhwy>>] [Accessed: 30/03/2011].
- Franklin, S. E., Maudie, A. J., and Lavigne, M. B., 2001. Using spatial co-occurrence texture to increase forest structure and species composition classification accuracy. *Photogrammetric Engineering and Remote Sensing*, 67 (7), 849-855.

- Frazer, G. W., Canham, C. D., and Lertzman, K. P., 1999. Gap light analyzer (gla): Imaging software to extract canopy structure and gap light transmission indices from true-colour fisheye photographs, users manual and program documentation. Burnaby, British Columbia, and Millbrook, New York: Simon Fraser University, Institute of Ecosystem Studies.
- Frazer, G. W., Fournier, R. A., Trofymow, J. A., and Hall, R. J., 2001. A comparison of digital and film fisheye photography for analysis of forest canopy structure and gap light transmission. *Agricultural and Forest Meteorology*, 109 (4), 249-263.
- Friedman, J., Hastie, T., and Tibshirani, R., 2013, Package 'glmnet'. [Online] [Available at: <<http://cran.r-project.org/web/packages/glmnet/glmnet.pdf>>] [Date accessed: 02/03/2013]
- Fuller, R., 2012. Avian responses to transitional habitats in temperate cultural landscapes: Woodland edges and young-growth. . In: Fuller, R. ed. *Birds and habitat. Relationships in changing landscapes*. Cambridge: Cambridge University Press, 125-149.
- Fuller, R. J., and Gill, R. M. A., 2001. Ecological impacts of increasing numbers of deer in british woodland. *Forestry*, 74 (3), 193-199.
- Gadow, K., 1999. Waldstruktur und diverzität. *AFJZ*, 170 (7), 117-121.
- Gamon, J. A., J. Penuelas, and Field, C.B. 1992. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment* 41, 35-44.
- Gamon, J.A. and J.S. Surfus, 1999. Assessing Leaf Pigment Content and Activity With a Reflectometer. *New Phytologist* 143, 105-117.
- Gaveau, L., and Hill, R., 2003. Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. *Canadian Journal of Remote Sensing*, 29 (5), 650-657.
- Geffen, E., Anderson, M. J., and Wayne, R. K., 2004. Climate and habitat barriers to dispersal in the highly mobile grey wolf. *Molecular Ecology*, 13 (8), 2481-2490.
- Gibbons, P., Ayers, D., Seddon, J., Doyle, S., and Briggs, S., 2004. Biometric decision support tool, version 1.4. *NSW department of environment and conservation, Canberra*, 51.
- Gill, S. J., Biging, G. S., and Murphy, E. C., 2000. Modeling conifer tree crown radius and estimating canopy cover. *Forest Ecology and Management*, 126 (3), 405-416.
- Gitelson, A. A., and Merzlyak, M. N., 1994. Spectral reflectance changes associated with autumn senescence of aesculus hippocastanum l. And acer platanoides l. Leaves. Spectral features and relation to chlorophyll estimation. *Journal of Plant Physiology* (143), 286-292. .
- Gitelson, A. A., Merzlyak, M. N., and Chivkunova, O. B., 2001. Optical properties and nondestructive estimation of anthocyanin content in plant leaves. *Photochemistry and Photobiology*, 74 (1), 38-45.
- Gitelson, A. A., Zur, Y., Chivkunova, O. B., and Merzlyak, M. N., 2002. Assessing carotenoid content in plant leaves with reflectance spectroscopy. *Photochemistry and Photobiology* 75, 272-281.

- Goetz, S., Steinberg, D., Dubayah, R., and Blair, B., 2007. Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. *Remote Sensing of Environment*, 108 (3), 254-263.
- Gougeon, F., and Leckie, D., 2003. *Forest information extraction from high spatial resolution images using an individual tree crown approach*. (Information report BC-X-396): Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre.
- Green, A. A., Berman, M., Switzer, P., and Craig, M. D., 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing*, 26 (1), 65-74.
- Gspaltl, M., Bauerle, W., Binkley, D., and Sterba, H., 2013. Leaf area and light use efficiency patterns of norway spruce under different thinning regimes and age classes. *Forest Ecology and Management*, 288, 49-59.
- Hall, S. A., Burke, I. C., Box, D. O., Kaufmann, M. R., and Stoker, J. M., 2005. Estimating stand structure using discrete-return lidar: An example from low density, fire prone ponderosa pine forests. *Forest Ecology and Management*, 208 (1-3), 189-209.
- Hastie, T., 2013. Generalized Additive Models. [Online] [Available from: <<http://cran.r-project.org/web/packages/gam/gam.pdf>>] [Date accessed: 24/04/2013]
- Hawthorne, L. B. 2012. Geospatial modelling environment. [Online][Available from: <<http://www.spatalecolology.com/gme>>] [Accessed: 16/08/2012].
- Heinzel, J., and Koch, B., 2011. Exploring full-waveform lidar parameters for tree species classification. *International Journal of Applied Earth Observation and Geoinformation*, 13 (1), 152-160.
- Heinzel, J., and Koch, B., 2013. Investigating multiple data sources for tree species classification in temperate forest and use for single tree delineation (vol 18, pg 101, 2012). *International Journal of Applied Earth Observation and Geoinformation*, 21, 581-581.
- Hendrickx, J., 2012. R package 'perturb' version 2.0.5. [Online] [Available from: <<http://cran.r-project.org/web/packages/perturb/>>] [Date accessed: 07/07/2012]
- Hendry, S. J., Boswell, R. C., and Rroudfoot, J. C. 2002. Forest condition 2001. In Commission, F. (Ed.), <http://www.forestry.gov.uk/>.
- Hill, R., Wilson, A., George, M., and Hinsley, S., 2009a. Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. *Journal of Applied Vegetation Science*.
- Hill, R. A., and Broughton, R. K., 2009b. Mapping the understorey of deciduous woodland from leaf-on and leaf-off airborne lidar data: A case study in lowland britain. *Isprs Journal of Photogrammetry and Remote Sensing*, 64 (2), 223-233.
- Hill, R. A., and Thomson, A. G., 2005. Mapping woodland species composition and structure using airborne spectral and lidar data. *International Journal of Remote Sensing*, 26 (17), 3763-3779.
- Hill, R. A., Wilson, A. K., George, M., and Hinsley, S. A., 2010. Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. *Applied Vegetation Science*, 13 (1), 86-99.

- Hinsley, S., Hill, R., Bellamy, P., and Balzter, H., 2006. The application of lidar in woodland bird ecology: Climate structure, and habitat quality. *Photogrammetric Engineering & Remote Sensing*, 72 (12), 1399-1406.
- Hinsley, S., Hill, R., Fuller, R., Bellamy, P., and Rothery, P., 2009a. Bird species distributions across woodland canopy structure gradients. *Community Ecology*, 10 (1), 99-110.
- Hinsley, S. A., Hill, R. A., Gaveau, D. L. A., and Bellamy, P. E., 2002. Quantifying woodland structure and habitat quality for birds using airborne laser scanning. *Functional Ecology*, 16 (6), 851-857.
- Höfle, B., and Pfeifer, N., 2007. Correction of laser scanning intensity data: Data and model-driven approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62 (6), 415-433.
- Holdridge, L. R., 1967. *Life zone ecology*. San Jose, Costa Rica: Tropical Science Centre.
- Holmgren, J., and Persson, A., 2004. Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*, 90 (4), 415-423.
- Holmgren, J., Persson, A., and Soderman, U., 2008. Species identification of individual trees by combining high resolution lidar data with multi-spectral images. *International Journal of Remote Sensing*, 29 (5), 1537-1552.
- Horner, G. J., Baker, P. J., Mac Nally, R., Cunningham, S. C., Thomson, J. R., and Hamilton, F., 2009. Mortality of developing floodplain forests subjected to a drying climate and water extraction. *Global Change Biology*, 15 (9), 2176-2186.
- Horner, G. J., Baker, P. J., Mac Nally, R., Cunningham, S. C., Thomson, J. R., and Hamilton, F., 2010. Forest structure, habitat and carbon benefits from thinning floodplain forests: Managing early stand density makes a difference. *Forest Ecology and Management*, 259 (3), 286-293.
- Huang, R., and He, M. Y., 2005. Band selection based on feature weighting for classification of hyperspectral data. *Ieee Geoscience and Remote Sensing Letters*, 2 (2), 156-159.
- Hudak, A. T., Crookston, N. L., Evans, J. S., Hall, D. E., and Falkowski, M. J., 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from lidar data. *Remote Sensing of Environment*, 112 (5), 2232-2245.
- Hudak, A. T., Evans, J. S., Crookston, N. L., Falkowski, M. J., Steigers, B. K., Taylor, R., and Hemingway, H., 2007, Feb 13-15. *Aggregating pixel-level basal area predictions derived from lidar data to industrial forest stands in north-central idaho*. Paper presented at the 3rd Forest Vegetation Simulator Conference, Ft Collins, CO.
- Hudak, A. T., Evans, J. S., and Smith, A. M. S., 2009. Lidar utility for natural resource managers. *Remote Sensing*, 1, 934-951.
- Hughes, G. F., 1968. On the mean accuracy of statistical pattern recognisers. *IEEE Trans. Inf. Theory*, 14 (1), 55-63.
- Hyde, P., Dubayah, R., Walker, W., Blair, J. B., Hofton, M., and Hunsaker, C., 2006. Mapping forest structure for wildlife habitat analysis using multi-sensor (lidar, sar/insar, etm plus , quickbird) synergy. *Remote Sensing of Environment*, 102 (1-2), 63-73.

- Hyypä, J., Hyypä, H., Leckie, D., Gougeon, F., Yu, X., and Maltamo, M., 2008. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing*, 29 (5), 1339-1366.
- Hyypä, J., Hyypä, H., Litkey, P., Yu, X., Haggren, H., Ronnholm, P., Pyysalo, J., Pitkanen, J., and Maltamo, M., 2004. Algorithms and methods of airborne laser-scanning for forest measurements. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 8/W2)* pp. 82–88.
- Hyypä, J., Kelle, O., Lehtikainen, M., and Inkinen, M., 2001. A segmentation-based method to retrieve stem volume estimates from 3-d tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*, 39 (5), 969-975.
- Isenburg, M., 2013. Lastools - efficient tools for lidar processing.[Online] [Available from: <<http://www.cs.unc.edu/~isenburg/lastools/>>] [Date accessed: 25/11/2012]
- JNCC. 2004. Common standards monitoring guidance for woodland habitats: Joint Nature Conservation Committee (GB). [Online] [Available from: <http://jncc.defra.gov.uk/pdf/CSM_woodland.pdf>] [Date accessed 09/09/2010]
- Jackson, R. D., Slater, P. N., and Pinter, P. J., 1983. Discrimination of growth and water stress in wheat by various vegetation indices through clear and turbid atmospheres. *Remote Sensing of the Environment* 15, 187-208.
- Jennings, S. B., Brown, N. D., and Sheil, D., 1999. Assessing forest canopies and understorey illumination: Canopy closure, canopy cover and other measures. *Forestry*, 72 (1), 59-73.
- Jensen, J. L. R., Humes, K. S., Vierling, L. A., and Hudak, A. T., 2008. Discrete return lidar-based prediction of leaf area index in two conifer forests. *Remote Sensing of Environment*, 112 (10), 3947-3957.
- Kaartinen, H., and Hyypä, J., 2008. Eurosd/isprrs project, commission ii “tree extraction”; finalreport; official publication no 53; eurosd (european spatial data research): Dublin, ireland.
- Kaartinen, H., Hyypä, J., Yu, X. W., Vastaranta, M., Hyypä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Naesset, E., Pitkanen, J., Popescu, S., Solberg, S., Wolf, B. M., and Wu, J. C., 2012. An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing*, 4 (4), 950-974.
- Kaasalainen, S., Hyypä, H., Kukko, A., Litkey, P., Ahokas, E., Hyypä, J., Lehner, H., Jaakkola, A., Suomalainen, J., Akujarvi, A., Kaasalainen, M., and Pyysalo, U., 2009. Radiometric calibration of lidar intensity with commercially available reference targets. *IEEE Transactions on Geoscience and Remote Sensing*, 47 (2), 588-598.
- Kalacska, M., Sanchez-Azofeifa, G. A., Calvo-Alvarado, J. C., Quesada, M., Rivard, B., and Janzen, D. H., 2004. Species composition, similarity and diversity in three successional stages of a seasonally dry tropical forest. *Forest Ecology and Management*, 200, 227-247.
- Kalliovirta, J., and Tokola, T., 2005. Functions for estimating stem diameter and tree age using tree height, crown width and existing stand database information. *Silva Fennica*, 39 (2), 227-248.

- Keddy, P. A., and Drummond, C. G., 1996. Ecological properties for the evaluation, management, and restoration of temperate deciduous forest ecosystems. *Ecological Applications*, 6 (3), 748-762.
- Ker, J. T., and Ostrovsky, M., 2003. From space to species: Ecological applications of remote sensing. *Trends in ecology and evolution*, 18, 299-305.
- Kim, S., Mcgaughey, R. J., Andersen, H. E., and Schreuder, G., 2009a. Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data. *Remote Sensing of Environment*, 113 (8), 1575-1586.
- Kim, Y., Yang, Z. Q., Cohen, W. B., Pflugmacher, D., Lauver, C. L., and Vankat, J. L., 2009b. Distinguishing between live and dead standing tree biomass on the north rim of grand canyon national park, USA using small-footprint lidar data. *Remote Sensing of Environment*, 113 (11), 2499-2510.
- Kint, V., Lust, N., Ferris, R., and Olsthoorn, A. F. M., 2000. Quantification of forest stand structure applied to scots pine (*pinus sylvestris* L.). *Forests. Invest. Agr. Sist. Recur. For.:Fuera de Serie.*, 1, 17p.
- Kirby, K. J., 2001. The impact of deer on the ground flora of british broadleaved woodland. *Forestry*, 74 (3), 219-229.
- Koop, H., Rijkssen, H. D., and Wind, J., 1994. Tools to diagnose forest integrity: An appraisal method substantiated by silvi-star assessment of biodiversity and forest structure. In: Boyle, T. J. B., Boontawee, B. (Eds.) ed. *Measuring and monitoring biodiversity in tropical and temperate forests*. Chaing Mai, Thailand: CIFOR, 309 – 331.
- Korpela, I., Hovi, A., and Morsdorf, F., 2012. Understory trees in airborne lidar data - selective mapping due to transmission losses and echo-triggering mechanisms. *Remote Sensing of Environment*, 119, 92-104.
- Koukoulas, S., and Blackburn, G. A., 2004. Quantifying the spatial properties of forest canopy gaps using lidar imagery and gis. *International Journal of Remote Sensing*, 25 (15), 3049-3071.
- Koukoulas, S., and Blackburn, G. A., 2005a. Mapping individual tree location, height and species in broadleaved deciduous forest using airborne lidar and multi-spectral remotely sensed data. *International Journal of Remote Sensing*, 26 (3), 431-455.
- Koukoulas, S., and Blackburn, G. A., 2005b. Spatial relationships between species and gap characteristics in broad-leaved deciduous woodland. *Journal of Vegetation Science*, 16, 587-597.
- Lada, H., Thomson, J. R., Mac Nally, R., Horrocks, G., and Taylor, A. C., 2007. Evaluating simultaneous impacts of three anthropogenic effects on a floodplain-dwelling marsupial antechinus flavipes. *Biological Conservation*, 134 (4), 527-536.
- Lähde, E., Laiho, O., Norokorpi, Y., and Saksa, T., 1999. Stand structure as the basis of diversity index. *Forest Ecology and Management*, 115 (2-3), 213-220.
- Langton, S. D., Briggs, P. A., and Haysom, K. A., 2010. Daubenton's bat distribution along rivers - developing and testing a predictive model. *Aquatic Conservation-Marine and Freshwater Ecosystems*, 20, S45-S54.
- Leckie, D., Gougeon, F., Hill, D., Quinn, R., Armstrong, L., and Shreenan, R., 2003a. Combined high-density lidar and multispectral imagery for individual tree crown analysis. *Canadian Journal of Remote Sensing*, 29 (5), 633-649.

- Leckie, D. G., Gougeon, F. A., Tims, S., Nelson, T., Burnett, C. N., and Paradine, D., 2005. Automated tree recognition in old growth conifer stands with high resolution digital imagery. *Remote Sensing of Environment*, 94 (3), 311-326.
- Leckie, D. G., Gougeon, F. A., Walsworth, N., and Paradine, D., 2003b. Stand delineation and composition estimation using semi-automated individual tree crown analysis. *Remote Sensing of Environment*, 85 (3), 355-369.
- Lee, A., Lucas, R., and Brack, C., 2004. Quantifying vertical forest stand structure using small footprint lidar to assess potential stand dynamics. In: *INTERNATIONAL ARCHIVES OF PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES*, Freiburg, Germany.
- Lee, A. C., and Lucas, R. M., 2007. A lidar-derived canopy density model for tree stem and crown mapping in australian forests. *Remote Sensing of Environment*, 111 (4), 493-518.
- Lefsky, M. A., Cohen, W. B., Parker, G. G., and Harding, D. J., 2002. Lidar remote sensing for ecosystem studies. *Bioscience*, 52 (1), 19-30.
- Lefsky, M. A., Cohen, W. B., and Spies, T. A., 2001. An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of douglas-fir forests in western oregon. *Canadian Journal of Forest Research*, 31 (1), 78-87.
- Leica Geosystems, 2003, ALS50 Airborne Laser Scanner technical specifications [Online] [Available from: <<http://www.leica-geosystems.com>>] [Date accessed: 24/07/2011]
- Lesak, A. A., Radeloff, V. C., Hawbaker, T. J., Pidgeon, A. M., Gobakken, T., and Contrucci, K., 2011. Modeling forest songbird species richness using lidar-derived measures of forest structure. *Remote Sensing of Environment*, 115 (11), 2823-2835.
- Levick, S. R. and Rogers, K.H. 2006. Lidar and object-based image analysis as tools for monitoring the structural diversity of savanna vegetation. In: *Int'l Archives for the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 34 (3), 6.
- Lexer, M. J., Lexer, W., and Hasenauer, H., 2000. The use of forest models for biodiversity assessment at the stand level. *Invest. Agr.: Sist. Recur. For.: Fuera de Serie*, 1, 297-316.
- Lim, K., Treitz, P., Baldwin, K., Morrison, I., and Green, J., 2003a. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. *Canadian Journal of Remote Sensing*, 29 (5), 658-678.
- Lim, K., Treitz, P., Wulder, M., St-Onge, B., and Flood, M., 2003b. Lidar remote sensing of forest structure. *Progress in Physical Geography*, 27 (1), 88-106.
- Lin, Y. C., and Mills, J. P., 2010. Factors influencing pulse width of small footprint, full waveform airborne laser scanning data. *Photogrammetric Engineering and Remote Sensing*, 76 (1), 49-59.
- Lindberg, E., Olofsson, K., Holmgren, J., and Olsson, H., 2012. Estimation of 3d vegetation structure from waveform and discrete return airborne laser scanning data. *Remote Sensing of Environment*, 118, 151-161.
- Lindenmayer, D. B., Franklin, J. F., and Fischer, J., 2006. General management principles and a checklist of strategies to guide forest biodiversity conservation. *Biological Conservation*, 131 (3), 433-445.

- Lovell, J. L., Jupp, D. L. B., Newnham, G. J., Coops, N. C., and Culvenor, D. S., 2005. Simulation study for finding optimal lidar acquisition parameters for forest height retrieval. *Forest Ecology and Management*, 214 (1-3), 398-412.
- Ltd., T. N., 2012. *Ecognition 8.7 object oriented image analysis user guide*. Munich.
- Lübbers. 1999. Diversitätsindizes und stichprobenverfahren. *Universität Freiburg*, 10p.
- Lucas, R., Bunting, P., Paterson, M., and Chisholm, L., 2008a. Classification of Australian forest communities using aerial photography, CASI and Hymap data. *Remote Sensing of Environment*, 112 (5), 2088-2103.
- Lucas, R., Lee, A., and Williams, M., 2006a. *The role of lidar data in understanding the relation between forest structure and SAR imagery*. Paper presented at the Proc. IGARSS 2005, Seoul, Korea, 25-29 Jul.
- Lucas, R. M., Cronin, N., Moghaddam, M., Lee, A., Armston, J., Bunting, P., and Witte, C., 2006b. Integration of radar and Landsat-derived foliage projected cover for woody regrowth mapping, Queensland, Australia. *Remote Sensing of Environment*, 100 (3), 388-406.
- Lucas, R. M., Lee, A. C., and Bunting, P. J., 2008b. Retrieving forest biomass through integration of CASI and lidar data. *International Journal of Remote Sensing*, 29 (5), 1553-1577.
- Lucas, R. M., Lee, A. C., and Williams, M. L., 2006c. Enhanced simulation of radar backscatter from forests using lidar and optical data. *IEEE Transactions on Geoscience and Remote Sensing*, 44 (10), 2736-2754.
- Lutz, D. A., Washington-Allen, R. A., and Shugart, H. H., 2008. Remote sensing of boreal forest biophysical and inventory parameters: A review. *Canadian Journal of Remote Sensing*, 34, S286-S313.
- MacArthur, R., and MacArthur, J. W., 1961. On bird species-diversity. *Ecology*, 42 (3), 594-598.
- MacDonald, E., and Hubert, J., 2002. A review of the effects of silviculture on timber quality of Sitka spruce. *Forestry*, 75 (2), 107-138.
- Magurran, A. E., 2004. *Measuring biological diversity*. Oxford: Blackwell.
- Mallet, C., Lafarge, F., Bretar, F., Soergel, U., Heipke, C., and Ieee. 2009. *Lidar waveform modeling using a marked point process*. New York: Ieee.
- Maltamo, M., Eerikainen, K., Pitkanen, J., Hyypä, J., and Vehmas, M., 2004. Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. *Remote Sensing of Environment*, 90 (3), 319-330.
- Maltamo, M., Packalen, P., Yu, X., Eerikainen, K., Hyypä, J., and Pitkanen, J., 2005. Identifying and quantifying structural characteristics of heterogeneous boreal forests using laser scanner data. *Forest Ecology and Management*, 216 (1-3), 41-50.
- Manes, F., Ricotta, C., Salvatori, E., Bajocco, S. and Blasi, C. (2010) A multiscale analysis of canopy structure in *Fagus sylvatica* L. and *Quercus cerris* L. old-growth forests in the Cilento and Vallo di Diano National Park. *Plant Biosystems*, 144 (1), 202-210.
- Martinuzzi, S., Vierling, L. A., Gould, W. A., Falkowski, M. J., Evans, J. S., Hudak, A. T., and Vierling, K. T., 2009. Mapping snags and understory shrubs for a lidar-based

- assessment of wildlife habitat suitability. *Remote Sensing of Environment*, 113 (12), 2533-2546.
- Mason, D. C., Anderson, G. Q. A., Bradbury, R. B., Cobby, D. M., Davenport, I. J., Vandepoll, M., and Wilson, J. D., 2003. Measurement of habitat predictor variables for organism-habitat models using remote sensing and image segmentation. *International Journal of Remote Sensing*, 24 (12), 2515-2532.
- Matthew, M. W., S.M. Adler-Golden, A. Berk, S.C. Richtsmeier, R.Y. Levine, L.S. Bernstein, P.K. Acharya, G.P. Anderson, G.W. Felde, M.P. Hoke, A. Ratkowski, H.-H. Burke, R.D. Kaiser, and Miller., D. P., 2000. Status of atmospheric correction using a modtran4-based algorithm. . *SPIE Proceeding, Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI*. 4049, pp. 199-207.
- Matthews, R., and Mackie, E., 2006. *Forest mensuration: A handbook for practitioners*. Forestry Commission, Edinburgh: HMSO.
- McDermid, G. J., Hall, R. J., Sanchez-Azofeifa, G. A., Franklin, S. E., Stenhouse, G. B., Kobliuk, T., and Ledrew, E. F., 2009. Remote sensing and forest inventory for wildlife habitat assessment. *Forest Ecology and Management*, 257 (11), 2262-2269.
- McElhinny, C., Gibbons, P., Brack, C., and Bauhus, J., 2005. Forest and woodland stand structural complexity: Its definition and measurement. *Forest Ecology and Management*, 218 (1-3), 1-24.
- McRoberts, R. E., Nelson, M. D., and Wendt, D. G., 2003. Stratified estimates of forest are using the k-nearest neighbours technique and satellite imagery. *General-Technical-Report-North-Centrral-research-Station, USDA-Forest-Service*, (NC-230), 80-86.
- Merganič, J., Merganičová, K., Marušák, R., and Audolenská, V. 2012. Plant diversity in forests. In Blanco, J. a. L., Y.(Ed) (Ed.), *Forest Ecosystems - More than Just Trees*
- Milton, E. J., Blackburn, G. A., Rollin, E. M., and Danson, F. M., 1994. Measurement of the spectral directional reflectance of forest canopies : A review of methods and a practical application. *Remote Sensing Reviews*, 10, 285-298.
- Miura, N., and Jones, S. D., 2010. Characterizing forest ecological structure using pulse types and heights of airborne laser scanning. *Remote Sensing of Environment*, 114 (5), 1069-1076.
- Moffiet, T., Mengersen, K., Witte, C., King, R., and Denham, R., 2005. Airborne laser scanning: Exploratory data analysis indicates potential variables for classification of individual trees or forest stands according to species. *Isprs Journal of Photogrammetry and Remote Sensing*, 59 (5), 289-309.
- Morsdorf, F., Kotz, B., Meier, E., Itten, K. I., and Allgower, B., 2006. Estimation of LAI and fractional cover from small footprint airborne laser scanning data based on gap fraction. *Remote Sensing of Environment*, 104 (1), 50-61.
- Mountford, E. P., Peterken, G. F., Edwards, P. J., and Manners, J. G., 1999. Long-term change in growth, mortality and regeneration of trees in denny wood, an old-growth wood-pasture in the new forest (uk). *Perspectives in Plant Ecology Evolution and Systematics*, 2 (2), 223-272.
- Mücke, W., Hollaus, M., and Pfeifer, N., 2012. Identification of dead trees using small footprint full-waveform airborne laser scanning data. In: *SilviLaser*, Vancouver, Canada.

- Mueller, T., Olson, K. A., Fuller, T. K., Schaller, G. B., Murray, M. G., and Leimgruber, P., 2008. In search of forage: Predicting dynamic habitats of mongolian gazelles using satellite-based estimates of vegetation productivity. *Journal of Applied Ecology*, 45 (2), 649-658.
- Murcia, C., 1995. Edge effects in fragmented forests - implications for conservation. *Trends in Ecology & Evolution*, 10 (2), 58-62.
- Mutlu, M., Popescu, S. C., Stripling, C., and Spencer, T., 2008. Mapping surface fuel models using lidar and multispectral data fusion for fire behavior. *Remote Sensing of Environment*, 112 (1), 274-285.
- Naesset, E., 1997a. Determination of mean tree height of forest stands using airborne laser scanner data. *Isprs Journal of Photogrammetry and Remote Sensing*, 52 (2), 49-56.
- Naesset, E., 1997b. Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment*, 61 (2), 246-253.
- Naesset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80 (1), 88-99.
- Naesset, E., 2004. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research*, 19 (2), 164-179.
- Naesset, E., Gobakken, T., Holmgren, J., Hyypä, H., Hyypä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, A., and Söderman, U., 2004. Laser scanning of forest resources: The nordic experience. *Scandinavian Journal of Forest Research*, 19 (6), 482-499.
- Naesset, E., and Okland, T., 2002. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment*, 79 (1), 105-115.
- NCAVEO, 2005, Geometric correction, [Online] [Available from: <http://www.ncaveo.ac.uk/geometric_correction/>] [Date accessed: 15/05/2012]
- Neumann, M., and Starlinger, F., 2001. The significance of different indices for stand structure and diversity in forests. *Forest Ecology and Management*, 145 (1-2), 91-106.
- Newton, A. C., Cantarello, E., Myers, G., Douglas, S., and Tejedor, N., 2010. The condition and dynamics of new forest woodlands. In: Newton, A. C. ed. *Biodiversity in the new forest*. Newbury, Berkshire: Pisces Publications, 132-147.
- Newton, A. C., Cayuela, L., Echeverria, C., Armesto, J. J., Del Castillo, R. F., Golicher, D., Geneletti, D., Gonzalez-Espinosa, M., Huth, A., Lopez-Barrera, F., Malizia, L., Manson, R., Premoli, A., Ramirez-Marcial, N., Benayas, J. M. R., Ruger, N., Smith-Ramirez, C., and Williams-Linera, G., 2009a. Toward integrated analysis of human impacts on forest biodiversity: Lessons from latin america. *Ecology and Society*, 14 (2), 41.
- Newton, A. C., Hill, R. A., Echeverria, C., Golicher, D., Benayas, J. M. R., Cayuela, L., and Hinsley, S. A., 2009b. Remote sensing and the future of landscape ecology. *Progress in Physical Geography*, 33 (4), 528-546.
- Newton, I., 1994. The role of nest sites in limiting the numbers of hole-nesting birds - a review. *Biological Conservation*, 70 (3), 265-276.
- Noss, R. F., 1990. Indicators for monitoring biodiversity - a hierarchical approach. *Conservation Biology*, 4 (4), 355-364.

- Noss, R. F., 1999. Assessing and monitoring forest biodiversity: A suggested framework and indicators. *Forest Ecology and Management*, 115 (2-3), 135-146.
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., and Wagner, H., 2013. R package 'vegan' version 2.0-2 [Online] [Available from: <<http://cran.r-project.org/web/packages/vegan/index.html>>] [Date accessed: 09/09/2012]
- Oldeland, J., Wesuls, D., Rocchini, D., Schmidt, M., and Jurgens, N., 2010. Does using species abundance data improve estimates of species diversity from remotely sensed spectral heterogeneity? *Ecological Indicators*, 10 (2), 390-396.
- Oliver, I., and Parkes, D., 2003. A prototype toolkit for scoring the biodiversity benefits of land use change. [Available from: <http://live.greeningaustralia.org.au/nativevegetation/pages/pdf/Authors%20O%20I%20Oliver_Parkes.pdf>] [Date accessed: 28/10/2010]
- Onojeghuo, A. O., and Blackburn, G. A., 2011. Optimising the use of hyperspectral and lidar data for mapping reedbed habitats. *Remote Sensing of Environment*, 115 (8), 2025-2034.
- Palmer, M. W., 1995. How should one count species. *Natural Areas Journal*, 15 (2), 124-135.
- Palmer, M. W., Earls, P. G., Hoagland, B. W., White, P. S., and Wohlgemuth, T., 2002. Quantitative tools for perfecting species lists. *Environmetrics*, 13 (2), 121-137.
- Parkes, D., Newell, G., and Cheal, D., 2003. Assessing the quality of native vegetation: The 'habitat hectares' approach. *Ecological Management and restoration*, 4, 29-38.
- Pascual, C., Garcia-Abril, A., Garcia-Montero, L. G., Martin-Fernandez, S., and Cohen, W. B., 2008. Object-based semi-automatic approach for forest structure characterization using lidar data in heterogeneous pinus sylvestris stands. *Forest Ecology and Management*, 255 (11), 3677-3685.
- Patenaude, G., Hill, R. A., Milne, R., Gaveau, D. L. A., Briggs, B. B. J., and Dawson, T. P., 2004. Quantifying forest above ground carbon content using lidar remote sensing. *Remote Sensing of Environment*, 93 (3), 368-380.
- Penuelas, J., F. Baret, and I. Filella, . 1995. Semi-empirical indices to assess carotenoids/chlorophyll-a ratio from leaf spectral reflectance. . *Photosynthetica* 31, 221-230. .
- Penuelas, J., Filella, I., Biel, C., Serrano, L., and Save, R., 1995. The reflectance at the 950-970 region as an indicator of plant water status. *International Journal of Remote Sensing* 14, 1887-1905.
- Persson, Å., Holmgren, J., and Soderman, U., 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing*, 68 (9), 925-932.
- Pesonen, Å., Maltamo, M., Eerikainen, K., and Packalen, P., 2008. Airborne laser scanning-based prediction of coarse woody debris volumes in a conservation area. *Forest Ecology and Management*, 255 (8-9), 3288-3296.
- Peterken, G., Spencer, J. W., and Field, A. P., 1996. *Maintaining the ancient and ornamental woods of the new forest*. Lyndhurst: Consultation document, Forestry Commission.

- Peterken, G. F., 2001. Ecological effects of introduced tree species in Britain. *Forest Ecology and Management*, 141 (1-2), 31-42.
- Pilot, M., Jedrzejewski, W., Branicki, W., Sidorovich, V. E., Jedrzejewska, B., Stachura, K., and Funk, S. M., 2006. Ecological factors influence population genetic structure of European grey wolves. *Molecular Ecology*, 15 (14), 4533-4553.
- Pike, R.J. and Wilson, S.E., 1971. Elevation relief ratio, hypsometric integral and geomorphic area altitude analysis. *Geological Society of America Bulletin*, 82, 1079-1084.
- Platt, W. J. and Strong, D. R. (eds.), 1989, Special feature: treefall gaps and forest dynamics. *Ecology*, 70, 535-576.
- Pommerening, A., 2002. Approaches to quantifying forest structures. *Forestry*, 75, 305-324.
- Pommerening, A., 2006. Evaluating structural indices by reversing forest structural analysis. *Forest Ecology and Management*, 224 (3), 266-277.
- Popescu, S. C., and Wynne, R. H., 2004. Seeing the trees in the forest: Using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. *Photogrammetric Engineering and Remote Sensing*, 70 (5), 589-604.
- Popescu, S. C., Wynne, R. H., and Nelson, R. F., 2003. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing*, 29 (5), 564-577.
- Popescu, S. C., Wynne, R. H., and Scrivani, J. A., 2004. Fusion of small-footprint lidar and multispectral data to estimate plot-level volume and biomass in deciduous and pine forests in Virginia, USA. *Forest Science*, 50 (4), 551-565.
- Popescu, S. C., and Zhao, K., 2008. A voxel-based lidar method for estimating crown base height for deciduous and pine trees. *Remote Sensing of Environment*, 112 (3), 767-781.
- Poulos, H. M., 2009. Mapping fuels in the Chihuahuan desert borderlands using remote sensing, geographic information systems, and biophysical modeling. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, 39 (10), 1917-1927.
- Pu, R. L., Kelly, M., Anderson, G. L., and Gong, P., 2008. Using CASI hyperspectral imagery to detect mortality and vegetation stress associated with a new hardwood forest disease. *Photogrammetric Engineering and Remote Sensing*, 74 (1), 65-75.
- Purves, D. W., Lichstein, J. W., and Pacala, S. W., 2007. Crown plasticity and competition for canopy space: A new spatially implicit model parameterized for 250 North American tree species. *Plos One*, 2 (9), 11.
- Putman, R., 2010. The effects of grazing on the ecological structure and dynamics of the new forest. In: Newton, A. C. ed. *Biodiversity in the new forest*. Oxford: Pisces Publications.
- Putman, R. J., 1996. Ungulates in temperate forest ecosystems: Perspectives and recommendations for future research. *Forest Ecology and Management*, 88 (1-2), 205-214.
- Quine, C. P., Cahalan, C., Hester, A., Humphrey, J., Kirby, K., Moffat, A., and Valatin, G., 2011. Woodlands chapter. In: Assessment, U. N. E. ed. *UK national ecosystem assessment - technical report*. Cambridge: UK National Ecosystem Assessment, UNEP-WCMC.

- Rautiainen, M., Heiskanen, J., Eklundh, L., Mottus, M., Lukes, P., and Stenberg, P., 2010. Ecological applications of physically based remote sensing methods. *Scandinavian Journal of Forest Research*, 25 (4), 325-339.
- Rautiainen, M., Suomalainen, J., Mottus, M., Stenberg, P., Voipio, P., Peltoniemi, J., and Manninen, T., 2007. Coupling forest canopy and understory reflectance in the arctic latitudes of finland. *Remote Sensing of Environment*, 110 (3), 332-343.
- Read, D. J., Freer-Smith, P. H., and Morison, J. I. L., 2009. *Combating climate change – a role for uk forests. The synthesis report*. The Stationary Office, Edinburgh.
- Reimoser, F., Armstrong, H., and Suchant, R., 1999. Measuring forest damage of ungulates: What should be considered. *Forest Ecology and Management*, 120 (1-3), 47-58.
- Reitberger, J., Krzystek, P., and Stilla, U., 2008. Analysis of full waveform lidar data for the classification of deciduous and coniferous trees. *International Journal of Remote Sensing*, 29 (5), 1407-1431.
- Reitberger, J., Schnorr, C., Krzystek, P., and Stilla, U., 2009. 3D segmentation of single trees exploiting full waveform lidar data. *Isprs Journal of Photogrammetry and Remote Sensing*, 64 (6), 561-574.
- Riano, D., Valladares, F., Condés, S. and Chuvueco, E., 2004. Estimation of leaf area index and covered ground from airborne laser scanner (LiDAR) in two contrasting forests. *Agricultural and Forest Meteorology*, 124, 269-275.
- Richards, J. A., 2012. *Remote sensing digital image analysis: An introduction (5th ed.)*. Berlin, Germany: Springer-Verlag.
- Richardson, A. D., and Berlyn, G. P., 2002. Changes in foliar spectral reflectance and chlorophyll fluorescence of four temperate species following branch cutting. *Tree Physiology*, 22 (7), 499-506.
- Richardson, J. J., and Moskal, L. M., 2011. Strengths and limitations of assessing forest density and spatial configuration with aerial lidar. *Remote Sensing of Environment*, 115 (10), 2640-2651.
- Rignot, E., Way, J., Williams, C., and Viereck, L., 1994. Radar estimates of aboveground biomass in boreal forests of interior alaska. *IEEE Transactions on Geoscience and Remote Sensing*, 32, 1117–1124.
- Rocchini, D., Balkenhol, N., Carter, G. A., Foody, G. M., Gillespie, T. W., He, K. S., Kark, S., Levin, N., Lucas, K., Luoto, M., Nagendra, H., Oldeland, J., Ricotta, C., Southworth, J., and Neteler, M., 2010a. Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges. *Ecological Informatics*, 5 (5), 318-329.
- Rocchini, D., He, K. S., Oldeland, J., Wesuls, D., and Neteler, M., 2010b. Spectral variation versus species beta-diversity at different spatial scales: A test in african highland savannas. *Journal of Environmental Monitoring*, 12 (4), 825-831.
- Rombouts, J., Ferguson, I. S., and Leech, J. W., 2008. *Variability of lidar volume prediction models for productivity assessment of radiata pine plantations in south australia*. Paper presented at the Proceedings of SilviLaser 2008: 8th international conference on LiDAR applications in forest assessment and inventory, Edinburgh, UK.

- Rosenqvist, A., Milne, A., Lucas, R., Imhoff, M., and Dobson, C., 2003. A review of remote sensing technology in support of the kyoto protocol. *Environmental Science & Policy*, 6 (5), 441-455.
- Rosette, J. a. B., North, P. R. J., and Suarez, J. C., 2008. Vegetation height estimates for a mixed temperate forest using satellite laser altimetry. *International Journal of Remote Sensing*, 29 (5), 1475-1493.
- Rouse, J. W., Haas, R. H., Schell, J. A., and Deering, D. W., 1973. Monitoring vegetation systems in the great plains with erts. In: *Third ERTS Symposium, NASA SP-351 I*: 309-317.
- Runkle, J. R., 1982. Patterns of disturbance in some old-growth mesic forests of eastern North America. *Ecology*, 63, 1533-1546.
- San, B.T. and Suzen, M. L., 2010, EVALUATION OF DIFFERENT ATMOSPHERIC CORRECTION ALGORITHMS FOREO-1 HYPERION IMAGERY. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Kyoto, Japan*, Volume 38, (8), 392-397.
- Sellers, P. J., 1985. Canopy reflectance, photosynthesis and transpiration. *International Journal of Remote Sensing* (6), 1335-1372.
- Serrano, L., Penuelas, J., and Ustin, S. L., 2002. Remote sensing of nitrogen and lignin in mediterranean vegetation from aviris data: Decomposing biochemical from structural signals. *Remote Sensing of Environment* 81 (355-364).
- Shannon, C. E., 1948. The mathematical theory of communication. In: Shannon, C. E., and Weaver, W. eds. *The mathematical theory of communication*. Urbana: University of Illinois Press, 29-125.
- Short, N.M. 2005. The Remote Sensing Tutorial [Online] [Available from: <http://www.fas.org/irp/imint/docs/rst/Sect13/Sect13_3.html>] [Date accessed: 15/03/2013]
- Simonson, W. D., Allen, H. D., and Coomes, D. A., 2013. Remotely sensed indicators of forest conservation status: Case study from a natura 2000 site in southern portugal. *Ecological Indicators*, 24, 636-647.
- Simpson, E. H., 1949. Measurement of diversity. *Nature*, 163, 688.
- Sims, D. A., and Gamon, J. A., 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment*, 81 (2-3), 337-354.
- Sithole, G., and Vosselman, G., 2004. Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. *Isprs Journal of Photogrammetry and Remote Sensing*, 59 (1-2), 85-101.
- Smith, A. M., Major, D. J., Hill, M. J., Willms, W. D., Brisco, B., and Lindwall, C. W., 1994. Airborne synthetic aperture radar analysis of rangeland revegetation of a mixed prairie. *Journal of Range Management*, 47, 385-391.
- Spies, T. A., F., J. F., and T., T. B., 1988. Coarse woody debris in douglas-fir forests of western oregon and washington. *Ecology*, 69, 1689-1702.
- St-Onge, B., Hu, Y., and Vega, C., 2008. Mapping the height and above-ground biomass of a mixed forest using lidar and stereo ikonos images. *International Journal of Remote Sensing*, 29 (5), 1277-1294.

- Staudhammer, C. L., and Lemay, V. M., 2001. Introduction and evaluation of possible indices of stand structural diversity. *Canadian Journal of Forest Research*, 31 (7), 1105-1115.
- Stokland, J. N., R., E., Tomter, S. M., Korhonen, K., Rajaniemi, S., Söderberg, U., Toet, H., and Riis Nielsen, T., 2003. *Forest biodiversity indicators in the nordic countries. Status based on national forest inventories.* . Copenhagen.
- Strunk, J., Temesgen, H., Andersen, H. E., Flewelling, J. P., and Madsen, L., 2012. Effects of lidar pulse density and sample size on a model-assisted approach to estimate forest inventory variables. *Canadian Journal of Remote Sensing*, 38 (5), 644-654.
- Suárez, J., 2004. *The use of aerial photography and ecognition analysis for tree counting in mature plantations. Project ordered by policy and practice division of the forestry commission and operational support unit from forest enterprise.*
- Thomas, V., Treitz, P., Mccaughey, J. H., Noland, T., and Rich, L., 2008. Canopy chlorophyll concentration estimation using hyperspectral and lidar data for a boreal mixedwood forest in northern ontario, canada. *International Journal of Remote Sensing*, 29 (4), 1029-1052.
- Tiede, D., Lang, S., and Hoffman, C., 2006. Supervised and forest type specific multi-scale segmentation for a one-level-representation of single trees. In: Lang, S., Blaschke, T., and Schöpfer, E. eds. *1st International Conference on Object-based Image Analysis (OBIA 2006)*. Salzburg University, Austria.
- Tits, L., De Keersmaecker, W., Somers, B., Asner, G. P., Farifteh, J., and Coppin, P., 2012. Hyperspectral shape-based unmixing to improve intra- and interclass variability for forest and agro-ecosystem monitoring. *Isprs Journal of Photogrammetry and Remote Sensing*, 74, 163-174.
- Tomppo, E., 1986. Models and methods for analysign spatial patterns of trees. . *Communicationses Institutu Forestalis Fenniae*, 138, 65p.
- Treitz, P., and Howarth, P., 1999a. Hyperspectral remote sensing for estimating biophysical parameters of forest ecosystems. *Progress in Physical Geography*, 23 (3), 359-390.
- Treitz, P., Lim, K., Woods, M., Pitt, D., Nesbitt, D., and Etheridge, D., 2012. Lidar sampling density for forest resource inventories in ontario, canada. *Remote Sensing*, 4 (4), 830-848.
- Treitz, P. M., and Howarth, P. J., 1999b. Hyperspectral remote sensing for estimating biophysical parameters of forest ecosystems. *Progress in Physical Geography*, 23 (3), 359-390.
- Tubbs, C. R., 2001. *The new forest. History, ecology, and conservation.* Lyndhurst, Hampshire: New Forest Ninth Centenary Trust
- Tucker, C. J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of the Environment*, 8, 127-150.
- Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., and Steininger, M., 2003. Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution*, 18 (6), 306-314.
- Van Den Meersschaut, D., and Vandekerckhove, K., 1998. Development of a stand scale forest biodiversity index based on the state forest inventory. In: Hansen, M., and Burk, T. eds. *Integrated tools for natural resources inventories in the 21st century.*

- General technical report nc-212*. Grove Boise, Idaho.: Proceedings held at Boise Centre August 16-20, 1998.
- Vaughn, N. R., Moskal, L. M., and Turnblom, E. C., 2012. Tree species detection accuracies using discrete point lidar and airborne waveform lidar. *Remote Sensing*, 4 (2), 377-403.
- Vauhkonen, J., Tokola, T., Maltamo, M., and Packalén, P., 2009. Applied 3d texture features in als-based forest inventory *European Journal of Forest Research*, 129, 803–811.
- Vellend, M., Verheyen, K., Flinn, K. M., Jacquemyn, H., Kolb, A., Van Calster, H., Peterken, G., Graae, B. J., Bellemare, J., Honnay, O., Brunet, J., Wulf, M., Gerhardt, F., and Hermy, M., 2007. Homogenization of forest plant communities and weakening of species-environment relationships via agricultural land use. *Journal of Ecology*, 95 (3), 565-573.
- Vierling, K. T., Vierling, L. A., Gould, W. A., Martinuzzi, S., and Clawges, R. M., 2008. Lidar: Shedding new light on habitat characterization and modeling. *Frontiers in Ecology and the Environment*, 6 (2), 90-98.
- Vogelmann, J. E., Rock, B. N., and Moss, D. M., 1993. Red edge spectral measurements from sugar maple leaves. . *International Journal of Remote Sensing* 14, 1563-1575.
- Vorčák, J., Merganic, J., and Saniga, M., 2006. Structural diversity change and regeneration processes of the norway spruce natural forest in babia hora nnr in relation to altitude. *Journal of Forest Science (Prague)*, 52 (9), 399-409.
- Wagner, W., Hyypä, J., Ullrich, A., Lehner, H., Briese, C., and Kaasalainen, S., 2008. Radiometric calibration of full-waveform small-footprint airborne laser scanners. *In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Beijing.
- Wagner, W., Ullrich, A., Ducic, V., Melzer, T., and Studnicka, N., 2006. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. *Isprs Journal of Photogrammetry and Remote Sensing*, 60 (2), 100-112.
- Wagner, W., Ullrich, A., Melzer, T., and K., K., 2004. From single-pulse to full-waveform airborne laser scanners: Potential and practical challenges. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35 (B3), 201-206.
- Wang, K., Franklin, S. E., Guo, X. L., and Cattet, M., 2010. Remote sensing of ecology, biodiversity and conservation: A review from the perspective of remote sensing specialists. *Sensors*, 10 (11), 9647-9667.
- Wang, L., Gong, P., and Biging, G. S., 2004. Individual tree-crown delineation and treetop detection high-spatial-resolution aerial imagery. *Photogrammetric Engineering and Remote Sensing*, 70 (3), 351-357.
- Wang, Y. S., Weinacker, H., and Koch, B., 2008. A lidar point cloud based procedure for vertical canopy structure analysis and 3d single tree modelling in forest. *Sensors*, 8 (6), 3938-3951.
- Waring, R. H., Way, J., Hunt Jr., E. R., Morrissey, L., Ranson, K. J., and Weishampel, J. F., 1995. Imaging radar for ecosystem studies. *BioScience*, 45, 715-723.

- Weiss, M., Baret, F., Smith, G. J., Jonckheere, I., and Coppin, P., 2004. Review of methods for in situ leaf area index next term (lai) determination: Part ii. Estimation of lai, errors and sampling. *Agricultural and Forest Meteorology*, 121 (1-2), 37-53.
- Woodhouse, L., 2006. *Introduction to microwave remote sensing*. New York: Taylor and Francis.
- Yu, X. W., Hyyppä, J., Kukko, A., Maltamo, M., and Kaartinen, H., 2006. Change detection techniques for canopy height growth measurements using airborne laser scanner data. *Photogrammetric Engineering and Remote Sensing*, 72 (12), 1339-1348.
- Zenner, E. K., 1999. Eine neue methode zur untersuchung der dreidimensionalität in waldbeständen. *Universität Freiburg*, 11p.
- Zenner, E. K., 2000. Do residual trees increase structural complexity in pacific northwest coniferous forests? *Ecological Applications*, 10, 800-810.
- Zerger, A., Gibbons, P., Seddon, J., Briggs, S., and Freudenberger, D., 2009. A method for predicting native vegetation condition at regional scales. *Landscape and Urban Planning*, 91 (2), 65-77.
- Zhang, J. K., Rivard, B., Sanchez-Azofeifa, A., and Castro-Esau, K., 2006. Intra and inter-class spectral variability of tropical tree species at la selva, costa rica: Implications for species identification using hydice imagery. *Remote Sensing of Environment*, 105 (2), 129-141.
- Zhang, K. Q., Chen, S. C., Whitman, D., Shyu, M. L., Yan, J. H., and Zhang, C. C., 2003. A progressive morphological filter for removing nonground measurements from airborne lidar data. *IEEE Transactions on Geoscience and Remote Sensing*, 41 (4), 872-882.
- Zheng, G., and Moskal, L. M., 2009. Retrieving leaf area index (lai) using remote sensing: Theories, methods and sensors. *Sensors*, 9 (4), 2719-2745.
- Zimble, D. A., Evans, D. L., Carlson, G. C., Parker, R. C., Grado, S. C., and Gerard, P. D., 2003. Characterizing vertical forest structure using small-footprint airborne lidar. *Remote Sensing of Environment*, 87 (2-3), 171-182.

Appendix A – Calculations applied to data collected in the field

A.1 Tree Populations and Regeneration

Total number of stems per hectare was estimated from the number recorded in the field plot, the plot area represented 9% of 1 hectare. Thus, it was possible to extrapolate up to this scale assuming similar conditions. This was also done for the number of native and total saplings and seedlings. Saplings were totalled and extrapolated to their occurrence at 1 hectare as with the trees. Seedlings were measured in the 10x10m sub-plot and thus represent 1% of 1 hectare.

A.2 Physical tree variables

Calculation of DBH from girth or circumference measurements was accomplished through the use of the formula for a circle:

$$DBH(m) = \frac{Girth(m)}{\pi}$$

Average DBH was calculated in addition to the standard deviation for each field plot. The percentage of big trees (DBH>0.4m) and very big trees (DBH>0.8m) relative to the tree population within the plot was calculated.

Similarly Basal Area (BA) for each tree was estimated with the formula for finding the area of a circle:

$$BA(m^2) = \pi \times \left(\frac{DBH(m)}{2}\right)^2$$

Field plot level BA (m²) is merely the sum of all of the individual trees. Plot estimates of BA correspond to 9% of the BA for 1 hectare.

A.3 Estimating tree height

Tree heights were derived through the use of trigonometry. Where assuming the tree went straight upward, its height was determined from the sum of two trigonometric equations, the 1st from the base of the tree to an angle of 0° (i.e. horizontal to the observer), and the 2nd from 0° to the top of the tree, both angular measurements were made at a set distance from the tree base, as illustrated in Figure A1. An example of the formulae is:

$$Tree\ Height = (A \times Tan(b)) + (A \times Tan(d))$$

Where the horizontal distance between the observer and the tree base is 'A', the angle below the horizontal (0°) to the base of the tree is 'b', the angle from horizontal to the tree top is 'd'.

Lorey's mean height is the mean height weighted by BA. This can be summarised as:

$$Lorey's\ Height\ (m) = \frac{\sum_i^n (DBH \times Height)}{Plot\ BA}$$

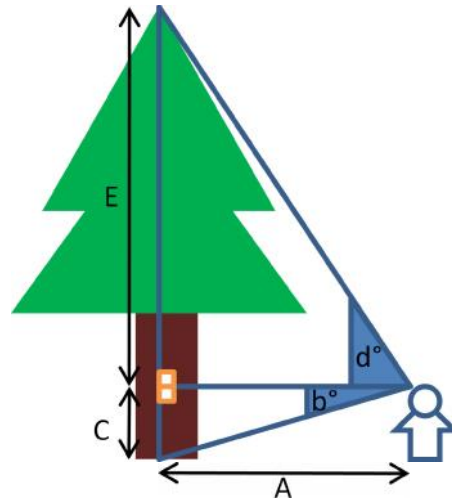


Figure A.1 – Estimation of tree height through the use of trigonometry. Where distance between observer and the tree base is ‘A’, the angle below the horizontal (0°) to the base of the tree is ‘b’, and the derived height to horizontal is ‘C’. The angle from horizontal to the tree top is ‘d’ and the derived height is ‘E’. Thus the sum of C and E represent the total tree height.

A.4 Tree crown variables

The height (m) of the crown base is calculated in the same manner as the total height of the tree in Figure A.1. The angle measured from the observer position to that of the first live branch is substituted for the angle to the tree top, in the previous equation.

Crown area was then estimated using the formula for an ellipse:

$$\text{Crown Area (m}^2\text{)} = \pi \times \left(\frac{\text{Crown Extent (N to S)}}{2} \right) \times \left(\frac{\text{Crown Extent (E to W)}}{2} \right)$$

A.5 Deadwood volume and variability

As in Cantarello and Newton (2008) the volume of downed deadwood (DDW) (both logs and branches) was determined using the equation for a frustum of a cone (i.e. a truncated cone), based on the individual measurements of log length and maximum and minimum diameters, illustrated in Figure A.2. The formula to find the volume of a frustum cone is:

$$Vol = \frac{\pi \times h}{3} \times (R^2 + R \times r + r^2)$$

When logs were circular in cross-section, medium diameter was measured and the equation for a cylinder was used instead. To reiterate, DDW measurements were limited to the 10x10m sub-plot. Volumes were then summed at the sub-plot level and transformed to m³ha⁻¹. DDW decay class was then averaged across the sub-plot.

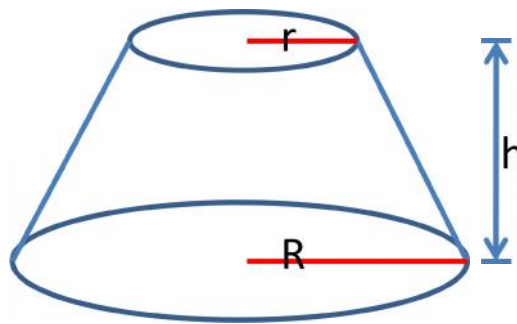


Figure A.2 – Frustum cone

Volume of snags was determined using the equation of either the frustum cone, or cylinder, dependent on trees' cross-section. Height/length were derived from clinometer and tape measurements. The volume was then summed at the plot level, and transformed to m³ha⁻¹. Snag decay class was also averaged for the whole plot.

A.6.Stem diameter differentiation index between neighbouring trees

Tree diameter differentiation has been identified as another potentially important parameter of structural diversity, for which an index was created by Földner (TM) (1995).

$$TM_i = \frac{1}{N} \sum_{j=1}^N (1 - d_{ij})$$

Where N is the number of trees on the sample plot, and d_{ij} is the relation between thinner and thicker DBH in the analysed neighbour tree pair. The index values range from 0 to 1.

However we are concerned with the sum of DBH differences between neighbouring trees. This can be expressed as:

$$sum\ diam.\ diff. = \sum_{j=1}^N (1 - d_{ij})$$

A.7 Index of vertical structure

Jaehne and Dohrenbusch (1997, in Newmann and Starlinger, 2001, Vorcak et al., 2006) proposed the Stand Diversity Index (B), the index of vertical structure. One of this index's component parts is the index of vertical structure.

$$S = 1 - \frac{\sum_{i=1}^N DBH_{min}}{\sum_{i=1}^N DBH_{max}}$$

Where N is the number of measured trees (3 thickest and 3 thinnest trees); DBH_{min} is the DBH of the thinnest trees (in cm); and DBH_{max} is the DBH of the thickest trees (in cm).

Appendix B– Hyperspectral data

B.1 – Spectral indices

Index:	Description:
Broadband Greenness – High pixel values indicate high proportions of green biomass	
Normalized Difference Vegetation Index (NDVI)	<p>The NDVI is one of the most frequently used vegetation indices. The index is robust over a wide range of conditions due to the use of the highest absorption and reflectance regions of Chlorophyll are used. It can however saturate in dense vegetation conditions, as LAI increases). The NDVI can be defined as:</p> $NDVI = \left(\frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} \right)$ <p>Where near infrared (<i>NIR</i>) and red bands (<i>red</i>) are used. Values can range between -1 and 1. The common range for green vegetation is 0.2 to 0.8.</p> <p>(Rouse et al., 1973; Tucker, 1979; Jackson et al., 1983; Sellers, 1985)</p>
Simple Ratio Index (SRI)	<p>The SR is the ratio of the highest reflectance. The use of the absorption wavelengths of chlorophyll makes it effective over many conditions. The SR can saturate in dense vegetation, i.e. when LAI becomes high. SR can be defined as:</p> $SR = \frac{\rho_{NIR}}{\rho_{red}}$ <p>The value of this index ranges from 0 to more than 30. The common range for green vegetation is 2 to 8.</p> <p>(Rouse et al., 1973; Tucker, 1979; Sellers, 1985)</p>
Atmospherically Resistant Vegetation Index (ARVI)	<p>The ARVI is an enhancement of the NDVI that is relatively resistant to atmospheric factors (e.g. aerosol). It uses reflectance in the blue wavelengths to correct the red reflectance for atmospheric scattering. The ARVI is defined as:</p> $ARVI = \frac{P_{NIR} - (2P_{RED} - P_{BLUE})}{P_{NIR} + (2P_{RED} - P_{BLUE})}$ <p>The value of this index ranges from -1 to 1. The common range for green vegetation is 0.2 to 0.8.</p>
Narrowband Greenness – High pixel values indicate high proportions of green biomass	
Red Edge Normalized Difference Vegetation Index (RENDVI)	<p>Is a modification of the standard broadband NDVI. This index differs by using bands along the red edge, instead of the main absorption and reflectance peaks. So the index will be sensitive to small changes in canopy foliage content, gap fraction and senescence. The inputs have been modified to sample features using bands centred on 705nm and 750nm. The index can be expressed as:</p> $RENDVI = \left(\frac{\rho_{705} - \rho_{750}}{\rho_{705} + \rho_{750}} \right)$ <p>The value of this index ranges from -1 to 1. The common range for green vegetation is 0.2 to 0.9.</p> <p>(Gitelson and Merzlyak, 1994; Sims and Gamon, 2002)</p>

B.1 – Spectral indices (continued)

Modified Red Edge Simple Ratio Index (MRESRI)	<p>The MRESRI is a modification of the broadband SR index. It uses bands in the red edge and incorporates a correction for leaf specular reflection. Applications include precision agriculture, forest monitoring, and vegetation stress detection. The index is defined by the following equation:</p> $MRESRI = \frac{P_{750} - P_{445}}{P_{705} - P_{445}}$ <p>The value of this index ranges from 0 to 30. The common range for green vegetation is 2 to 8.</p> <p>(Sims and Gamon, 2002; Datt, 1999)</p>
Modified Red Edge Normalized Difference Vegetation Index (MRENDVI)	<p>The MRENDVI is a modification of the Red Edge NDVI. It incorporates a correction for leaf specular reflection. The index capitalizes on the sensitivity of the vegetation red edge to small changes in canopy foliage content, gap fraction, and senescence. Applications include precision agriculture, forest monitoring, and vegetation stress detection. The index is defined by the following equation:</p> $MRENDVI = \frac{P_{750} - P_{705}}{P_{750} + P_{705} - 2P_{445}}$ <p>The value of this index ranges from -1 to 1. The common range for green vegetation is 0.2 to 0.7.</p> <p>(Datt, 1999; Sims and Gamon, 2002)</p>
Vogelmann Red Edge Index 1 (VOG1)	<p>(<i>Narrowband</i>) The VOG1 is sensitive to the effects of foliage chlorophyll concentration, canopy leaf area, and water content. The index is defined by the equation:</p> $VOG1 = \frac{\rho_{740}}{\rho_{720}}$ <p>The value of this index ranges from 0 to 20. The common range for green vegetation is 4 to 8.</p> <p>(Vogelmann, et al., 1993)</p>
Light Use Efficiency – Will measure of the efficiency with which vegetation is able to use incident light for photosynthesis.	
Photochemical Reflectance Index (PRI)	<p>The PRI is a measure of the reflectance which is sensitive to changes in carotenoid pigments (mainly xanthophylls) in live foliage. Carotenoid pigments are related to the efficiency of photosynthetic light use, or the uptake rate of carbon per unit of energy absorbed. It can be used to study vegetation productivity and stress. PRI is defined as:</p> $PRI = \frac{\rho_{531} - \rho_{570}}{\rho_{531} + \rho_{570}}$ <p>The value of this index ranges from -1 to 1. The common range for green vegetation is -0.2 to 0.2.</p> <p>(Gamon et al., 1992; 1997)</p>

B.1 – Spectral indices (continued)

Structure Insensitive Pigment Index (SIPI)	<p>The SIPI is designed to maximize the sensitivity of the index to the ratio of bulk carotenoids to chlorophyll while decreasing sensitivity to variation in canopy structure (for example, leaf area index). Increases in SIPI are thought to indicate increased canopy stress (carotenoid pigment). SIPI is defined as:</p> $SIPI = \frac{\rho_{800} - \rho_{445}}{\rho_{800} + \rho_{680}}$ <p>The value of this index ranges from 0 to 2. The common range for green vegetation is 0.8 to 1.8.</p> <p>(Penuelas et al., 1995)</p>
Red Green Ratio Index (RGR)	<p>The RGR indicates the relative expression of leaf redness caused by anthocyanin to that of chlorophyll. The RG Ratio has been used to estimate the course of foliage development in canopies. The RG Ratio index is an indicator of leaf production and stress, and may also indicate flowering in some canopies.</p> <p>The value of this index ranges from 0.1 to more than 8. The common range for green vegetation is 0.7 to 3.</p> <p>(Gamon and Surfus, 1999)</p>
Leaf Pigments – are designed to provide a measure of stress-related pigments present in vegetation.	
Carotenoid Reflectance Index 1 (CRI1)	<p>The CRI1 is sensitive to carotenoid pigments in plant foliage. Higher CRI1 values mean greater carotenoid concentration relative to chlorophyll. CRI1 is defined by the following equation:</p> $CRI1 = \left(\frac{1}{\rho_{510}} \right) - \left(\frac{1}{\rho_{550}} \right)$ <p>The value of this index ranges from 0 to more than 15. The common range for green vegetation is 1 to 12.</p> <p>(Gitelson et al., 2002)</p>
Anthocyanin Reflectance Index 1 (ARI1)	<p>ARI1 is sensitive to anthocyanins in plant foliage. Increases in ARI1 indicate canopy changes in foliage via new growth or death. ARI1 is defined by the following equation:</p> $ARI1 = \left(\frac{1}{\rho_{550}} \right) - \left(\frac{1}{\rho_{700}} \right)$ <p>The value of this index ranges from 0 to more than 0.2. The common range for green vegetation is 0.001 to 0.1.</p> <p>(Gitelson et al., 2001)</p>
Canopy Water Content – applications include canopy stress analysis, productivity prediction and modelling, fire hazard condition analysis, cropland management, and studies of ecosystem physiology	
Water Band Index (WBI)	<p>The WBI is sensitive to changes in canopy water status. As the water content of vegetation canopies increases, the strength of the absorption around 970 nm increases relative to that of 900 nm. WBI is defined by the following equation:</p> $WBI = \frac{\rho_{900}}{\rho_{970}}$ <p>The common range for green vegetation is 0.8 to 1.2.</p> <p>(Penuelas et al., 1995; Champagne et al., 2001)</p>

B.2 Hyperspectral metric list– A full listing of all extracted hyperspectral area-based metrics from the eCognition derived species classification map and ENIV derived spectral indices input into statistical modelling. (See appendix C.2 for index descriptions.)

Metric Name:	Description:	Statistics included:
“no_tree”	Number of tree crown polygons per 30x30m area.	[n=1] No. tree crown polygons
“no_nat_t”	Number of tree crown polygons per 30x30m area of native species.	[n=1] No. native species tree crown polygons
“spec_no”	Number of tree species detected per 30x30m area.	[n=1] No. detected tree species
“nat_spec”	Number of tree species detected per 30x30m area of native species	[n=1] No. native detected tree species
“SH”	Shannon Index calculated from tree crown polygons per 30x30m area.	[n=1] Shannon index value
“SI”	Simpson Index calculated from tree crown polygons per 30x30m area.	[n=1] Simpson index value
“ndvil...”	NDVI index summarised to the 30x30m area. Leaf-off (1) and leaf-on (2) data.	[n=6] Minimum, maximum, range, mean, standard deviation and sum.
“sri1...”	SRI index summarised to the 30x30m area. Leaf-off (1) and leaf-on (2) data.	[n=6] Minimum, maximum, range, mean, standard deviation and sum.
“arvil...”	ARVI index summarised to the 30x30m area. Leaf-off (1) and leaf-on (2) data.	[n=6] Minimum, maximum, range, mean, standard deviation and sum.
“rendvil...”	RENDVI index summarised to the 30x30m area. Leaf-off (1) and leaf-on (2) data.	[n=6] Minimum, maximum, range, mean, standard deviation and sum.
“mresril...”	MRESRI index summarised to the 30x30m area. Leaf-off (1) and leaf-on (2) data.	[n=6] Minimum, maximum, range, mean, standard deviation and sum.
“mrendvil...”	MRENDVI index summarised to the 30x30m area. Leaf-off (1) and leaf-on (2) data.	[n=6] Minimum, maximum, range, mean, standard deviation and sum.
“vreil...”	VREI index summarised to the 30x30m area. Leaf-off (1) and leaf-on (2) data.	[n=6] Minimum, maximum, range, mean, standard deviation and sum.
“sipi1...”	SIPI index summarised to the 30x30m area. Leaf-off (1) and leaf-on (2) data.	[n=6] Minimum, maximum, range, mean, standard deviation and sum.
“rgri1...”	RGRI index summarised to the 30x30m area. Leaf-off (1) and leaf-on (2) data.	[n=6] Minimum, maximum, range, mean, standard deviation and sum.
“pri1...”	PRI index summarised to the 30x30m area. Leaf-off (1) and leaf-on (2) data.	[n=6] Minimum, maximum, range, mean, standard deviation and sum.
“ari1...”	ARI index summarised to the 30x30m area. Leaf-off (1) and leaf-on (2) data.	[n=6] Minimum, maximum, range, mean, standard deviation and sum.
“cri1...”	CRI index summarised to the 30x30m area. Leaf-off (1) ONLY.	[n=6] Minimum, maximum, range, mean, standard deviation and sum.
“wbi1...”	WBI index summarised to the 30x30m area. Leaf-off (1) and leaf-on (2) data.	[n=6] Minimum, maximum, range, mean, standard deviation and sum.

B.3 eCognition classification summary R script – R code developed to extract/calculate hyperspectral area-based metrics from eCognition derived species classification map.

Overview: Custom R code was developed to estimate area based metrics relating to overstorey tree species for each grid cell (30x30m area) iteratively. Firstly, each of the ITC objects within a grid cell (Cell ID) subset was counted. Using subset functions, counts of native species only was extracted. Number of species encountered for each Cell ID, and native species was counted. Finally, estimates of Shannon, Simpson and Evenness indices were applied to the input species list. For each Cell ID an estimate of tree count, native tree count, species count, native species count, maximum number of individuals (from one species group), minimum number of individuals (from one species group) Shannon, Simpson and Evenness indices were computed.

```
#####Library foreign#####
library(foreign)#needed for reading and writing to dbf format
#####Library Vegan#####
library(Vegan)#needed for Shannon and Simpson index calculation
#Load data
tree.data<- read.dbf("Classificaton_Output.dbf")
#
#create results table#
results<- array(NA,c(0,7))
#
###select by polygon ID###
#Enter minimum ID no.
ID<-3033
#loop until max. ID no.
while (ID <= 44030) {
print(ID)
#subset from dataset the rows with the current polygon ID no.
tree.subset<-tree.data[which(tree.data$CELLID==ID),]
#
###ITC object count###
class<-tree.subset$Class_name
class<-as.data.frame(table(class))
no.tree<-sum(class$Freq)
#
###Count the no. native ITC objects(species list: Oak, Beech, Silver Birch, Scots Pine,
Common Adler)###
c_OK<-class[which(class=='Oak'),]
c_BE<-class[which(class=='Beech'),]
c_SB<-class[which(class=='Silver Birch'),]
c_SP<-class[which(class=='Scots Pine'),]
c_CA<-class[which(class=='Common_Adler'),]
no_n_tree<-sum(c_OK$Freq, c_BE$Freq, c_SB$Freq, c_SP$Freq, c_CA$Freq)
#
###ITC species count###
species<-tree.subset$Class_name
species<-as.data.frame(table(unique(species)))
spec_no<-sum(species$Freq)
#
###Count the number of native species (species list: Oak, Beech, Silver Birch, Scots Pine,
Common Adler)###
```



```

#species
s_OK<-species[which(class=='Oak'),]
s_BE<-species[which(class=='Beech'),]
s_SB<-species[which(class=='Silver Birch'),]
s_SP<-species[which(class=='Scots Pine'),]
s_CA<-species[which(class=='Common Adler'),]
no_n_spec<-sum(s_OK$Freq, s_BE$Freq, s_SB$Freq, s_SP$Freq, s_CA$Freq)
#
####Calculate Shannon Index###
Shannon <- diversity(class$Freq, "shannon")
Simpson <- diversity(class$Freq, "simpson")
#
#CalculateEvenness index#
Evenness<-Shannon/log2(spec_no)
#
####CSDI-part 1 inputs: max/min pop###
#count of each class within polygon
class_a<- class$Freq
#removes spec.freq with 0 values
class_a<-class_a[which(class_a!=0)]
#
#check if pop has more than two tree spec.
if (spec_no>1) {
  max_pop<- max(class_a)      #maximum species count value returned
  min_pop<- min(class_a)      #minimum species count value returned
}
else
{
  max_pop<- 0
  min_pop<- 0
}
####output row###
IMS<-c(ID,no.tree, no_n_tree, spec_no, no_n_spec, Shannon, Simpson, Evenness, Max_no,
Min_no)
IMS<-as.data.frame(t(IMS))
ID<-ID+1                      #ID vaule plus 1
#
####write to output table###
#
results<-merge(IMS,results,all=TRUE)
#
rm(tree.subset,class,no.tree,no_n_tree,species,spec_no,no_n_spec,Shannon,Simpson,Evenness,
s , IMS,max_pop, min_pop)
#
}
colnames(results)<-c("ID", "no_tree", "no_nat_t","spec_no","nat_spec","SH","SI","Even",
"max_pop", "min_pop")
write.csv(results,file="Class_metrics.csv", row.names=T)
write.dbf(results, file="Class_metrics2.dbf")

```

Appendix C –LiDAR data

C.1 Extracted metrics from DR LiDAR for 30x30 grid cells

Metric:	Metric description:	Software:									
SHA_mean	Mean height of all returns	RSC Las tools									
SHA_Var	Variance of height for all returns	RSC Las tools									
SHA_skw	Skewness of height for all returns.	RSC Las tools									
SHA_kur	Kurtosis of height for all returns.	RSC Las tools									
SHA_max	Maximum height recorded	RSC Las tools									
SHA_med	Median height for all returns	RSC Las tools									
SHA_STD	Standard deviation of heights for all returns	RSC Las tools									
SHA_MAD	Mean Absolute Deviation of heights for all returns	RSC Las tools									
SHA_Range	Range of height values for all returns	RSC Las tools									
SHN_mean	Mean height of all non-ground returns	RSC Las tools									
SHN_Var	Variance of height for non-ground returns	RSC Las tools									
SHN_skw	Skewness of height for non-ground returns	RSC Las tools									
SHN_kur	Kurtosis of height for non-ground returns	RSC Las tools									
SHN_max	Maximum height recorded for non-ground	RSC Las tools									
SHN_med	Median height for non-ground returns	RSC Las tools									
SHN_STD	Standard deviation of heights for non-ground returns	RSC Las tools									
SHN_MAD	Mean Absolute Deviation of heights for non-ground returns	RSC Las tools									
SHN_Range	Range of height values for non-ground returns	RSC Las tools									
DTM_slope	<p>The topographic slope estimated from a DEM generated from the ground classified points (Burrough and McDonell, 1998). This metric is commonly expressed as a raster.</p> $\text{slope_degrees} = \text{ATAN}(\text{rise_run}) * 57.29578$ <p>where:</p> $\text{rise_run} = \sqrt{[dz/dx]^2 + [dz/dy]^2}$ <p>The values of the centre cell and its eight neighbours determine the horizontal and vertical deltas. The neighbours are identified as letters from 'a' to 'i', with 'e' representing the cell for which the aspect is being calculated.</p> <table border="1"> <tr> <td>a</td><td>b</td><td>c</td></tr> <tr> <td>d</td><td>e</td><td>f</td></tr> <tr> <td>g</td><td>h</td><td>i</td></tr> </table> <p>The rate of change in the x direction for cell 'e' is calculated with the algorithm:</p> $[dz/dx] = ((c + 2f + i) - (a + 2d + g)) / (8 * x_cell_size)$ <p>The rate of change in the y direction for cell 'e' is calculated with the following algorithm:</p> $[dz/dy] = ((g + 2h + i) - (a + 2b + c)) / (8 * y_cell_size)$	a	b	c	d	e	f	g	h	i	RSC Las tools
a	b	c									
d	e	f									
g	h	i									
DTM_rough	The topographic roughness estimated from a DEM generated from the ground classified points. This is commonly expressed as a raster. The calculation is based upon the standard deviation of recorded elevation in relation to its neighbouring points for the centre value in a 3x3 raster cell grid.	RSC Las tools									
CC	<p>Canopy cover – as a percentage. Calculated from the ratio of ground vs. non-ground returns. This is defined by the equation:</p> $CC = \left(\frac{h_{ng}}{h_{all}} \right)$ <p>Where h_{ng} and h_{all} denote the sum total of non-ground returns and the sum of all returns.</p>	RSC Las tools									

C.1(continued)

Frac_cov	<p>Fractional cover – estimating the proportion of an area that is covered by vegetation above 0.5m. As defined by Morsdorf et al (2006) the equation:</p> $fCover = \frac{\sum E_{vegetation}}{\sum E_{total}}$ <p>With</p> $E_{vegetation} = E_{total} > 0.5m$ <p>Where $E_{vegetation}$ and E_{total} denotes vegetation returns and all (ground and vegetation) returns respectively.</p>	RSC Las tools
CR_ratio	<p>Canopy Relief ratio – As a measure of the relative shape of the canopy from LiDAR observations the elevation relief ratio, E (Pike and Wilson, 1971), was utilised. This is defined as:</p> $E = \frac{h_{mean} - h_{min}}{h_{max} - h_{min}}$ <p>Where h_{mean}, h_{min} and h_{max} are the mean, minimum and maximum canopy heights, respectively. This ratio reflects the degree to which outer canopy surfaces are in the upper ($E > 0.5$) or in the lower ($E < 0.5$) portions of the height range.</p>	RSC Las tools
PAIP	<p>(For leaf-on data only) The Plant Area Index Proxy – similar to the LAI, however this includes all vegetative components (leaves and woody stems) (Morsdorf et al, 2006). The equation:</p> $PAIP = \frac{\sum E_{FR}}{\sum E_{LR} + E_{SR}}$ <p>E_{FR}, E_{SR} and E_{LR} denote the tree types of returns, the first return, the last return and the single return (first return = last return), but only for vegetation. The vegetation points are classified by thresholding the height over the terrain, as with fractional cover this is 0.5m</p>	RSC Las tools
P005	Percentage counts of LiDAR values located in the 5 th percentile of canopy height	RSC Las tools
P010	Percentage counts of LiDAR values located in the 10 th percentile of canopy height	RSC Las tools
P015	Percentage counts of LiDAR values located in the 15 th percentile of canopy height	RSC Las tools
P020	Percentage counts of LiDAR values located in the 20 th percentile of canopy height	RSC Las tools
P025	Percentage counts of LiDAR values located in the 25 th percentile of canopy height	RSC Las tools
P030	Percentage counts of LiDAR values located in the 30 th percentile of canopy height	RSC Las tools
P035	Percentage counts of LiDAR values located in the 35 th percentile of canopy height	RSC Las tools
P040	Percentage counts of LiDAR values located in the 40 th percentile of canopy height	RSC Las tools
P045	Percentage counts of LiDAR values located in the 45 th percentile of canopy height	RSC Las tools
P050	Percentage counts of LiDAR values located in the 50 th percentile of canopy height	RSC Las tools
P055	Percentage counts of LiDAR values located in the 55 th percentile of canopy height	RSC Las tools
P060	Percentage counts of LiDAR values located in the 60 th percentile of canopy height	RSC Las tools
P065	Percentage counts of LiDAR values located in the 65 th percentile of canopy height	RSC Las tools

C.1(continued)

P070	Percentage counts of LiDAR values located in the 70 th percentile of canopy height	RSC Las tools
P075	Percentage counts of LiDAR values located in the 75 th percentile of canopy height	RSC Las tools
P080	Percentage counts of LiDAR values located in the 80 th percentile of canopy height	RSC Las tools
P085	Percentage counts of LiDAR values located in the 85 th percentile of canopy height	RSC Las tools
P090	Percentage counts of LiDAR values located in the 90 th percentile of canopy height	RSC Las tools
P095	Percentage counts of LiDAR values located in the 95 th percentile of canopy height	RSC Las tools
can_layer	No. Canopy Layers	SPDlib
can_depth	Canopy Depth	SPDlib
max_gap	Canopy maximum gap	SPDlib

C.2 Extracted intensity related metrics from DR LiDAR data for 30x30 grid cells(RSC LASTools)

Metric:	Metric description:	Software:
SIA_mean	Mean intensity of all returns	RSC Las tools
SIA_Var	Variance of intensity for all returns	RSC Las tools
SIA_skw	Skewness of intensity for all returns	RSC Las tools
SIA_kur	Kurtosis of intensity for all returns	RSC Las tools
SIA_max	Maximum intensity recorded	RSC Las tools
SIA_med	Median intensity for all returns	RSC Las tools
SIA_STD	Standard deviation of intensity for all returns	RSC Las tools
SIA_MAD	Mean Absolute Deviation of intensity for all returns	RSC Las tools
SIA_Range	Range of intensity values for all returns	RSC Las tools
SIG_mean	Mean intensity of ground returns	RSC Las tools
SIG_Var	Variance of intensity for ground returns	RSC Las tools
SIG_skw	Skewness of intensity for ground returns	RSC Las tools
SIG_kur	Kurtosis of intensity for ground returns	RSC Las tools
SIG_max	Maximum intensity recorded for ground returns	RSC Las tools
SIG_med	Median intensity for ground returns	RSC Las tools
SIG_STD	Standard deviation of intensity for ground returns	RSC Las tools
SIG_MAD	Mean Absolute Deviation of intensity for ground returns	RSC Las tools
SIG_Range	Range of intensity values for ground returns	RSC Las tools
SIN_mean	Mean intensity of non-ground returns	RSC Las tools
SIN_Var	Variance of intensity for ground returns	RSC Las tools
SIN_skw	Skewness of intensity for ground returns	RSC Las tools
SIN_kur	Kurtosis of intensity for ground returns	RSC Las tools
SIN_max	Maximum intensity recorded for ground returns	RSC Las tools
SIN_med	Median intensity for ground returns	RSC Las tools
SIN_STD	Standard deviation of intensity for ground returns	RSC Las tools
SIN_MAD	Mean Absolute Deviation of intensity for ground returns	RSC Las tools
SIN_Range	Range of intensity values for ground returns	RSC Las tools
p005i	The LiDAR intensity classified as the 5 th percentile of canopy height	RSC Las tools
p010i	The LiDAR intensity classified as the 10 th percentile of canopy height	RSC Las tools
p015i	The LiDAR intensity classified as the 15 th percentile of canopy height	RSC Las tools
p020i	The LiDAR intensity classified as the 20 th percentile of canopy height	RSC Las tools
p025i	The LiDAR intensity classified as the 25 th percentile of canopy height	RSC Las tools
p030i	The LiDAR intensity classified as the 30 th percentile of canopy height	RSC Las tools
p035i	The LiDAR intensity classified as the 35 th percentile of canopy height	RSC Las tools
p040i	The LiDAR intensity classified as the 40 th percentile of canopy height	RSC Las tools
p045i	The LiDAR intensity classified as the 45 th percentile of canopy height	RSC Las tools
p050i	The LiDAR intensity classified as the 50 th percentile of canopy height	RSC Las tools
p055i	The LiDAR intensity classified as the 55 th percentile of canopy height	RSC Las tools
p060i	The LiDAR intensity classified as the 60 th percentile of canopy height	RSC Las tools
p065i	The LiDAR intensity classified as the 65 th percentile of canopy height	RSC Las tools
p070i	The LiDAR intensity classified as the 70 th percentile of canopy height	RSC Las tools
p075i	The LiDAR intensity classified as the 75 th percentile of canopy height	RSC Las tools
p080i	The LiDAR intensity classified as the 80 th percentile of canopy height	RSC Las tools
p085i	The LiDAR intensity classified as the 85 th percentile of canopy height	RSC Las tools
p090i	The LiDAR intensity classified as the 90 th percentile of canopy height	RSC Las tools
p095i	The LiDAR intensity classified as the 95 th percentile of canopy height	RSC Las tools

C.3 Derived TIFFS outputs for 30x30 grid cells

Metric:	Metric description:	Calculation:
TIF_TreeNo	A count of all tree crown centres located within a 30x30m grid area	Calculated externally in R*
TIF_space	A mean distance between tree crown centre points	Calculated externally in R*
TIF_STD	A standard deviation between tree crown centre points	Calculated externally in R*
TIF_CR	A mean of all canopy radii for each of the tree crown centres encountered in the 30x30m grid area.	Calculated externally in R*
TIF_m_area	Each tree crown point had the crown area calculated using the crown radius. The mean of all crown areas was made.	Calculated externally in R*
TIF_to_area	Each tree crown point had the crown area calculated using the crown radius. The total of all crown areas was made.	Calculated externally in R*
TIF_HT	The mean of all tree heights within a 30x30m grid area (m).	Calculated externally in R*
TIF_meanCV	The mean geometric crown volume within the 30x30m grid area.	Calculated externally in R*
TIF_totalCV	The total geometric crown volume within the 30x30m grid area.	Calculated externally in R*

*See appendix C.4

C.4R script for summarising ITC metrics– Annotated R code developed to calculate the average distance and standard deviation of points within a 30x30m polygon cell.

Overview: the tree list data has gone through a spatial join operation in ArcMap, identifying which trees tree points intersected with which grid-polygon. The code iteratively cycles through each polygon ID and identifies which trees are present within, and then applies distance measurements from each point to every other point. An average distance and standard deviation is then calculated and summary metrics calculated.

```
library(foreign)#needed for reading and writing to dbf format
#
###load data###
tree.data<- read.dbf("TIFFS_points.dbf")
#create results table#
results<- array(NA,c(0,10))
#
###Distance function###
f<-function(w) apply(xy,1,function(v) sum((v-w)^2))
###select lowest polygon ID no. to begin###
ID<-1
# loop until max ID no.#
while (ID <= 22) {
  print(ID)
  #subset dataset by polygon ID
  tree.subset<-tree.data[which(tree.data$Site_no==ID),]
  #error check – metrics cannot be calculated with 1 ITC point
  if (nrow(tree.subset)>1) {
    #Enter X and Y cords
    x<-tree.subset$POINT_X
    y<-tree.subset$POINT_Y
    #
    ###compute all distances between distinct points###
    xy<-cbind(x,y)
    distances2<-apply(xy,1,f)
    distances2<-sqrt(distances2)
    #Tree/ITC space metrics require minimum distances
    minOfColumns=apply(distances2, 2, function(x) min(x[x!=0]) )
    TIF_Space<- mean(minOfColumns)
    TIF_STD <- sd(minOfColumns)
    #
    #ITC attributes summaries
    TIF_TreeNo<- nrow(tree.subset)#number of points
    TIF_HT <- mean(tree.subset$treeHt)#average point height
    TIF_CR <- mean(tree.subset$crownRad)#average crown radius
    TIF_m_area<- mean(pi*tree.subset$crownRad^2)#average crown horizontal area
    TIF_to_area<- sum(pi*tree.subset$crownRad^2)#total crown horizontal area
    TIF_meanCV<- mean(tree.subset$canopyVol)#average crown geometric volume
    TIF_totalCV<- sum(tree.subset$canopyVol)#total crown geometric volume
    #
    IMS<-c(ID, TIF_TreeNo, TIF_Space, TIF_STD, TIF_CR, TIF_m_area, TIF_to_area,
    TIF_HT, TIF_meanCV, TIF_totalCV) #output row
```



```

IMS<-as.data.frame(t(IMS))
ID<-ID+1          #ID value plus 1
#
###write to output table###
#
results<-merge(IMS,results,all=TRUE)
#
#remove inputs so it doesn't influence future loops.
rm(x,y,xy,distances2,TIF_TreeNo, TIF_Space, TIF_STD, TIF_CR, TIF_m_area,
TIF_to_area, TIF_HT, TIF_meanCV, TIF_totalCV, IMS)
}
else
{
#####
###if no points detected skip to next ID###
ID<-ID+1          #ID value plus 1
#
}
#####
}
#end loop#
#rename columns
colnames(results)<-c("ID", "TIF_TreeNo", "TIF_Space", "TIF_STD", "TIF_CR",
"TIF_m_area", "TIF_to_area", "TIF_HT", "TIF_meanCV", "TIF_totalCV")
#output as both csv and dbf format...
write.csv(results,file="TIFFS_metrics_f2012.csv", row.names=T)
write.dbf(results, file="TIFFS_metrics2_f2012.dbf")

```

C.5 ITC metric extraction for CSDI– Annotated R code developed to calculate the three smallest and three largest distances of trees within the 30x30m polygon cell, in addition to the two smallest and 2 largest horizontal crown diameters.

Overview: the tree list data has gone through a spatial join operation in ArcMap, identifying which trees tree points intersected with which grid-polygon. The code iteratively cycles through each polygon ID and identifies which trees are present within, and then applies distance measurements from each point to every other point. The three smallest and three largest distances of trees within the 30x30m polygon cell, in addition to the two smallest and 2 largest horizontal crown diameters are extracted and output. These metrics are necessary for the Complex stand diversity index (Jaehne and Dohrenbusch, 1997 in Newmann and Starlinger, 2001).

```
####Library=foreign####
library(foreign)
#
####load data####
tree.data<- read.dbf("Tiffs_pnts_XY.dbf")#to open .dbf format files
#create results table#
results<- array(NA,c(0,11)) #11 columns for 10 outputs + ID
#
####Distance function####
f<-function(w) apply(xy,1,function(v) sum((v-w)^2))
####select by polygon lowest polygon ID no.
ID<-1
#begin loop until max poly ID no. reached
while (ID <= 21) {
  print(ID)
  tree.subset<-tree.data[which(tree.data$CELLID==ID),]
  if (nrow(tree.subset)>5) {      #Requires a minimum of 6 trees to compute
    x<-tree.subset$POINT_X
    y<-tree.subset$POINT_Y
    xy<-cbind(x,y)
    #
    ####compute all distances between distinct points####
    distances2<-apply(xy,1,f)
    distances2<-sqrt(distances2)
    #output minimum distances between ITC points
    minOfColumns=apply(distances2, 2, function(x) min(x[x!=0]) )
    #
    ####spatial distribution (3 biggest and 3 smallest)####
    ####3 minimum distances####
    minOfColumns<-sort(minOfColumns)#sort so min. at top
    Min.d1<-min(minOfColumns)      #1st smallest
    Min.d2<-minOfColumns[[2]]      #2nd smallest
    Min.d3<-minOfColumns[[3]]      #3rd smallest
    #
    ####3 maximum distances####
    #re-order distances so maximum is at top
    minOfColumns<-sort(minOfColumns,decreasing = TRUE)
    Max.d1<-max(minOfColumns)      #1st biggest
    Max.d2<-minOfColumns[[2]]      #2nd biggest
    Max.d3<-minOfColumns[[2]]      #3rd biggest
  }
```

```

#
###crown differentiation (2 biggest and 2 smallest diameters)###
tree.subset2<-tree.subset$crownRad
tree.subset2<-2*tree.subset2      #convert radius to diameter
#
#Min crowns#
tree.subset2<-sort(tree.subset2)#sort so min. at top
Min.cr1<-min(tree.subset2)  #1st smallest
Min.cr2<-tree.subset2[[2]]  #2nd smallest
#Max crowns#
#re-orders distances so maximum is at top
tree.subset2<-sort(tree.subset2,decreasing = TRUE)
Max.cr1<-max(tree.subset2)  #1st biggest
Max.cr2<-tree.subset2[[2]]  #2nd biggest
#
#output row#
IMS<-
c(ID,Min.d1,Min.d2,Min.d3,Max.d1,Max.d2,Max.d3,Min.cr1,Min.cr2,Max.cr1,Max.cr2)
IMS<-as.data.frame(t(IMS))
ID<-ID+1      #ID value plus 1
#
###write to output table###
results<-merge(IMS,results,all=TRUE)
#
rm(x,y,xy,distances2,Min.d1,Min.d2,Min.d3,Max.d1,Max.d2,Max.d3,Min.cr1,Min.cr2,
Max.cr1,Max.cr2,IMS)
}
Else      #if tree population is ≤1 skip to next ID
{
#####
ID<-ID+1      #ID value plus 1
#
}
#####
}
#end loop#
#Re-name results columns
colnames(results)<-
c("ID","Min.d1","Min.d2","Min.d3","Max.d1","Max.d2","Max.d3","Min.cr1","Min.cr2","Ma
x.cr1","Max.cr2")
#output csv and dbf files
write.csv(results,file="TIFFS_CSDI_part.csv", row.names=T)
write.dbf(results, file="TIFFS_CSDI_part_2.dbf")

```

C.6 Extracted metrics from FWLiDAR data for 30x30 grid cells

Metric:	Metric description:	Software:
CC	Canopy cover – as a percentage. Calculated from the ratio of ground vs. non-ground returns. This is defined by the equation: $CC = \left(\frac{h_{ng}}{h_{all}} \right)$ Where h_{ng} and h_{all} denote the sum total of non-ground returns and the sum of all returns.	SPDlib
Meanht	Mean height of all returns	SPDlib
Meannght	Mean height of all non-ground returns	SPDlib
Medht	Median height of all returns	SPDlib
Mednght	Median height of all non-ground returns	SPDlib
Maxht	Maximum point height recorded	SPDlib
Domht	The dominant height – average height of the non-ground first returns in the highest 20% of returns.	SPDlib
Stdht	Standard deviation of all heights	SPDlib
stdnght	Standard deviation of all non-ground heights	SPDlib
Varht	Variance of all height values	SPDlib
Varnght	Variance of non-ground height values	SPDlib
Adevht	The absolute deviation of all height values	SPDlib
Adevnght	The absolute deviation of non-ground height values	SPDlib
Skwht	Skewness of all height values	SPDlib
Skwnght	Skewness of non-ground height values	SPDlib
Kurht	Kurtosis of all height values	SPDlib
Kurnght	Kurtosis of non-ground height values	SPDlib
Ht_p05	Percentage counts of LiDAR values located in the 5 th percentile of canopy height	SPDlib
Ht_p10	Percentage counts of LiDAR values located in the 10 th percentile of canopy height	SPDlib
Ht_p15	Percentage counts of LiDAR values located in the 15 th percentile of canopy height	SPDlib
Ht_p20	Percentage counts of LiDAR values located in the 20 th percentile of canopy height	SPDlib
Ht_p25	Percentage counts of LiDAR values located in the 25 th percentile of canopy height	SPDlib
Ht_p30	Percentage counts of LiDAR values located in the 30 th percentile of canopy height	SPDlib
Ht_p35	Percentage counts of LiDAR values located in the 35 th percentile of canopy height	SPDlib
Ht_p40	Percentage counts of LiDAR values located in the 40 th percentile of canopy height	SPDlib
Ht_p45	Percentage counts of LiDAR values located in the 45 th percentile of canopy height	SPDlib
Ht_p50	Percentage counts of LiDAR values located in the 50 th percentile of canopy height	SPDlib
Ht_p55	Percentage counts of LiDAR values located in the 55 th percentile of canopy height	SPDlib
Ht_p60	Percentage counts of LiDAR values located in the 60 th percentile of canopy height	SPDlib
Ht_p65	Percentage counts of LiDAR values located in the 65 th percentile of canopy height	SPDlib
Ht_p70	Percentage counts of LiDAR values located in the 70 th percentile of canopy height	SPDlib
Ht_p75	Percentage counts of LiDAR values located in the 75 th percentile of canopy height	SPDlib
Ht_p80	Percentage counts of LiDAR values located in the 80 th percentile of canopy height	SPDlib
Ht_p85	Percentage counts of LiDAR values located in the 85 th percentile of canopy height	SPDlib
Ht_p90	Percentage counts of LiDAR values located in the 90 th percentile of canopy height	SPDlib
Ht_p95	Percentage counts of LiDAR values located in the 95 th percentile of canopy height	SPDlib
can_layer	No. Canopy Layers	SPDlib
can_depth	Canopy Depth	SPDlib
max_gap	Canopy maximum gap	SPDlib

C.7 – Extracted metrics from FW LiDAR data for 30x30 grid cells (amplitude)

Metric:	Metric Description:	Software:
Meanamp	Mean amplitude value for all returns	SPDlib
Meangamp	Mean amplitude value for ground returns	SPDlib
Meanngamp	Mean amplitude value for non-ground returns	SPDlib
Medamp	Mean amplitude value for all returns	SPDlib
Medgamp	Mean amplitude value for ground returns	SPDlib
Medngamp	Mean amplitude value for non-ground returns	SPDlib
Minamp	Minimum amplitude value for all returns	SPDlib
Mingamp	Minimum amplitude value for ground returns	SPDlib
Minngamp	Minimum amplitude value for non-ground returns	SPDlib
Maxamp	Maximum amplitude value for all returns	SPDlib
Maxgamp	Maximum amplitude value for ground returns	SPDlib
Maxngamp	Maximum amplitude value for non-ground returns	SPDlib
Stdamp	Standard deviation of amplitude value for all returns	SPDlib
Stdgamp	Standard deviation of amplitude value for ground returns	SPDlib
Stdngamp	Standard deviation of amplitude value for non-ground returns	SPDlib
Varamp	Variance of amplitude value for all returns	SPDlib
Vargamp	Variance of amplitude value for ground returns	SPDlib
Varngamp	Variance of amplitude value for non-ground returns	SPDlib
Adevamp	Absolute deviation of amplitude value for all returns	SPDlib
Adevgamp	Absolute deviation of amplitude value for ground returns	SPDlib
Adevngamp	Absolute deviation of amplitude value for non-ground returns	SPDlib
Skwamp	Skewness of amplitude value for all returns	SPDlib
Skwgamp	Skewness of amplitude value for ground returns	SPDlib
Skwngamp	Skewness of amplitude value for non-ground returns	SPDlib
Kuramp	Kurtosis of amplitude value for all returns	SPDlib
Kurgamp	Kurtosis of amplitude value for ground returns	SPDlib
Kurngamp	Kurtosis of amplitude value for non-ground returns	SPDlib
Amp_p05	The LiDAR intensity classified as the 5 th percentile of canopy height	SPDlib
Amp_p10	The LiDAR intensity classified as the 10 th percentile of canopy height	SPDlib
Amp_p15	The LiDAR intensity classified as the 15 th percentile of canopy height	SPDlib
Amp_p20	The LiDAR intensity classified as the 20 th percentile of canopy height	SPDlib
Amp_p25	The LiDAR intensity classified as the 25 th percentile of canopy height	SPDlib
Amp_p30	The LiDAR intensity classified as the 30 th percentile of canopy height	SPDlib
Amp_p35	The LiDAR intensity classified as the 35 th percentile of canopy height	SPDlib
Amp_p40	The LiDAR intensity classified as the 40 th percentile of canopy height	SPDlib
Amp_p45	The LiDAR intensity classified as the 45 th percentile of canopy height	SPDlib
Amp_p50	The LiDAR intensity classified as the 50 th percentile of canopy height	SPDlib
Amp_p55	The LiDAR intensity classified as the 55 th percentile of canopy height	SPDlib
Amp_p60	The LiDAR intensity classified as the 60 th percentile of canopy height	SPDlib
Amp_p65	The LiDAR intensity classified as the 65 th percentile of canopy height	SPDlib
Amp_p70	The LiDAR intensity classified as the 70 th percentile of canopy height	SPDlib
Amp_p75	The LiDAR intensity classified as the 75 th percentile of canopy height	SPDlib
Amp_p80	The LiDAR intensity classified as the 80 th percentile of canopy height	SPDlib
Amp_p85	The LiDAR intensity classified as the 85 th percentile of canopy height	SPDlib
Amp_p90	The LiDAR intensity classified as the 90 th percentile of canopy height	SPDlib
Amp_p95	The LiDAR intensity classified as the 95 th percentile of canopy height	SPDlib

C.8 – Extracted metrics from FW LiDAR data for 30x30 grid cells (width)

Metric:	Metric Description:	Software:
Meanwidth	Mean width value for all returns	SPDlib
Meangwidth	Mean width value for ground returns	SPDlib
Meangamp	Mean width value for non-ground returns	SPDlib
Medwidth	Mean width value for all returns	SPDlib
Medgwidth	Mean width value for ground returns	SPDlib
Medngwidth	Mean width value for non-ground returns	SPDlib
Minwidth	Minimum width value for all returns	SPDlib
Mingwidth	Minimum width value for ground returns	SPDlib
Minngwidth	Minimum width value for non-ground returns	SPDlib
Maxwidth	Maximum width value for all returns	SPDlib
Maxgwidth	Maximum width value for ground returns	SPDlib
Maxngwidth	Maximum width value for non-ground returns	SPDlib
Stdwidth	Standard deviation of width value for all returns	SPDlib
Stdgwidth	Standard deviation of width value for ground returns	SPDlib
Stdngwidth	Standard deviation of width value for non-ground returns	SPDlib
Varwidth	Variance of width value for all returns	SPDlib
Vargwidth	Variance of width value for ground returns	SPDlib
Varngwidth	Variance of width value for non-ground returns	SPDlib
Adevwidth	Absolute deviation of width value for all returns	SPDlib
Adevgwidth	Absolute deviation of width value for ground returns	SPDlib
Adevngwidth	Absolute deviation of width value for non-ground returns	SPDlib
Skwidth	Skewness of width value for all returns	SPDlib
Skwgwidth	Skewness of width value for ground returns	SPDlib
Skwngwidth	Skewness of width value for non-ground returns	SPDlib
Kurwidth	Kurtosis of width value for all returns	SPDlib
Kurgwidth	Kurtosis of width value for ground returns	SPDlib
Kurngwidth	Kurtosis of width value for non-ground returns	SPDlib
Wd_p05	The LiDAR width classified as the 5 th percentile of canopy height	SPDlib
Wd_p10	The LiDAR width classified as the 10 th percentile of canopy height	SPDlib
Wd_p15	The LiDAR width classified as the 15 th percentile of canopy height	SPDlib
Wd_p20	The LiDAR width classified as the 20 th percentile of canopy height	SPDlib
Wd_p25	The LiDAR width classified as the 25 th percentile of canopy height	SPDlib
Wd_p30	The LiDAR width classified as the 30 th percentile of canopy height	SPDlib
Wd_p35	The LiDAR width classified as the 35 th percentile of canopy height	SPDlib
Wd_p40	The LiDAR width classified as the 40 th percentile of canopy height	SPDlib
Wd_p45	The LiDAR width classified as the 45 th percentile of canopy height	SPDlib
Wd_p50	The LiDAR width classified as the 50 th percentile of canopy height	SPDlib
Wd_p55	The LiDAR width classified as the 55 th percentile of canopy height	SPDlib
Wd_p60	The LiDAR width classified as the 60 th percentile of canopy height	SPDlib
Wd_p65	The LiDAR width classified as the 65 th percentile of canopy height	SPDlib
Wd_p70	The LiDAR width classified as the 70 th percentile of canopy height	SPDlib
Wd_p75	The LiDAR width classified as the 75 th percentile of canopy height	SPDlib
Wd_p80	The LiDAR width classified as the 80 th percentile of canopy height	SPDlib
Wd_p85	The LiDAR width classified as the 85 th percentile of canopy height	SPDlib
Wd_p90	The LiDAR width classified as the 90 th percentile of canopy height	SPDlib
Wd_p95	The LiDAR width classified as the 95 th percentile of canopy height	SPDlib

C.9 – Description of additional FW attributes

Amplitude

The echo amplitude recorded by LiDAR systems is commonly referred to as ‘intensity’ despite the fact that in physical terms it would be more natural to associate the intensity with the total energy of one echo, whilst the amplitude measurement only characterises the peak power of the echo (Wagner et al., 2008). Here we define backscatter as the reflection of the laser pulses back in the direction of the receiver. It is a diffuse reflection due to scattering, as opposed to specular reflection like a mirror.

According to Alexander et al (2010) the amplitude values for a given object would vary depending upon flying height or elevation differences from the emitted laser pulse, even within a single dataset. There would be a requirement for the conversion or correction of amplitude values from multi-temporal analysis, different airborne LiDAR systems, scan geometry and atmospheric conditions so comparisons could be made. The additional information provided by the decomposition the full backscatter waveform is suitable for calibration, where amplitude and echo width are converted into values proportional to the surface reflectance of the target object (Hofle and Pfeifer, 2007; Kaasalainen et al., 2009).

It should be noted however that echo amplitude and waveform measurements, such as those identified above, depend not only on the backscattering properties of the target objects but also on sensor and flight parameters, such as flying altitude, beam divergence (Alexander et al., 2010), footprint size (Li, 2008), laser pulse energy, atmospheric conditions, etc. Wagner et al (2008) states that the amplitude and waveform measurements from different sensor designs, acquisition campaigns, and flight strips are not directly comparable. It may not be possible to compare the measurements taken within one flight strip because of topographic height variations and variable atmospheric conditions along the flight path. It is therefore necessary to perform some kind of relative correction, or more desirably to convert the echo amplitude and echo measurements into physical parameters describing the backscatter properties in a quantitative way (Wagner et al., 2008).

Pulse Width

As mentioned previously backscatter waveforms are popularly modelled by a mathematical function such as the Gaussian function. The extracted parameter representing pulse width can be used to evaluate the extent of pulse broadening. The pulse width could refer to either the Full Width at Half Maximum (FWHM) amplitude, or the standard deviation, of the echo in the Gaussian decomposition (Alexander et al., 2010). A number of studies on small-footprint waveform systems can be used to improve range accuracy and multi-target resolution, especially in complex waveforms (Chauve et al., 2007). In relation to the ability to extract additional data from the backscatter waveform, it is believed that they can potentially improve classification of data points. The capability of not only backscatter cross-section, amplitude, but also pulse width and the number of returns within the waveforms have been investigated for the classification of land cover objects (Lin and Mills, 2010).

Pulse width has been demonstrated as having an important role in 3D vegetation mapping (Reitberger et al., 2008). According to Lin and Mills (2010) review a number of studies related to large-footprint LiDAR systems, where three main factors cause pulse-broadening: surface roughness, nadir angle, and surface slope. A factorial analysis was then applied, showing that roughness was the most influential factor affecting pulse response. Slope and scan angle have little, but similar effects upon the pulse width. They conclude that when compared with conventional laser intensity values, pulse width can there be considered as a

universal and relatively stable parameter for small footprint systems to provide information about surface characteristics. Caution must be applied however in the cases of interpreting pulse widths from areas to tree canopy and weak pulses (e.g. the forest floor).

The classification of features, such ground points, can be expected using pulse width information to identify vegetation in open areas. In areas where multiple returns are encountered, such as over forest canopy, it may be difficult to classify areas based solely on pulse width (Lin and Mills, 2010). This is due to the micro-scale surface information present within the small footprint size available, for example 0.2 of 2m. Therefore returns from tree branches or buildings can exhibit behaviour of smooth surfaces. Also, estimates for pulse width from weak returns, possibly near the forest floor, can cause erroneous values. Additional criteria, for example the spatial relationships between points can be incorporated as a measure to reduce noise.

Conversely, Adams et al. (2012) research indicates that FW parameters relating to amplitude, peak-width, and exponential decay constant varied substantially across all surface types, ruling out the potential to determine source characteristics for individual returns for forest attributes. The exponential decay constant is the return shape between the peak and the next local minima. However it was observed that pulses on the ground on average had a great intensity, decay constant and a narrower peak than returns from coniferous foliage.

C.10 Vertical evenness index calculated from ITC R script– R code developed to estimate the Vertical Evenness index (Neumann and Starlinger, 2001) from TIFFS point data.

Overview: Custom R code was developed to estimate the Vertical Evenness index (Neumann and Starlinger, 2001) from TIFFS point data. The code goes through each of the necessary calculations required in construction of the modified Shannon Index for each grid cell (30x30m area) iteratively. Firstly, each of the trees within the 30x30m area is assigned a height class (1-4) based on its relative height to the maximum tree height within the plot. The rows are then subset by the assigned classification, and the total horizontal areas summed at the subset level. Index calculations are then applied, e.g. applying logarithms, and a final index (VE) value is output for each cell ID.

```
#####Library foreign#####
library(foreign)#so dbf files can be read-in
#
tree.data<- read.dbf("Tiffs_pnts.dbf")#Tiffs derived tree crown centroid locations
#
#create results table#
results<- array(NA,c(0,2))#output of 2 columns
#
#####select by polygon ID###
ID<-1
#begin loop#
while (ID <= 21) {
  print(ID)
  tree.subset<-tree.data[which(tree.data$CELLID==ID),]
  if (nrow(tree.subset)>1) {#clause for if there is only 1 tree in the 30x30m cell#
  #
  #####subset by height group: (1)0-19%, (2)20-49%, (3)50-79%, and (4)80-100%###
  max<-max(tree.subset$treeHt) #max height
  #
  req<-cbind(tree.subset$treeHt, tree.subset$crown_area)
  #
  req_2 <- apply(req,1,function(x) (x[1]/max)*100) #determine percentage of max. Height#
  req_3<-cbind(req,req_2)
  #
  #Classification Function#
  a<-function(x) ifelse (req_3[,3]<=19, "1", ifelse (req_3[,3]<=49, "2", ifelse (req_3[,3]<=79,
  "3", ifelse (req_3[,3]<=100, "4", "NA"))))
  #
  req_4<-apply(req_3, 1, a)
  req_5<-cbind(req_3,req_4[,1]) #height group set
  colnames(req_5)<-c("HT", "Area", "per","HG") #renames column headings
  req_5<-as.data.frame(req_5)
  #
  #####Total Area Calc###
  AA<-req_5$Area
  AA<-as.character(AA)
  AA<-as.numeric(AA)
  A_TA<-sum(AA)
  #
  #####subset by height groups - calculate relative area###
  #1(0-19)
```

```

group.1<-req_5[ which(req_5[,4]=='1'),]
group.1a<-group.1$Area
group.1a<-as.character(group.1a) #numeric error work-around
group.1a<-as.numeric(group.1a)
g1_TA<-sum(group.1a)
#
#2(20-49)
group.2<-req_5[ which(req_5[,4]=='2'),]
group.2a<-group.2$Area
group.2a<-as.character(group.2a)
group.2a<-as.numeric(group.2a)
g2_TA<-sum(group.2a)
#
#3 (50-79)
group.3<-req_5[ which(req_5[,4]=='3'),]
group.3a<-group.3$Area
group.3a<-as.character(group.3a)
group.3a<-as.numeric(group.3a)
g3_TA<-sum(group.3a)
#
#4 (>80)
group.4<-req_5[ which(req_5[,4]=='4'),]
group.4a<-group.4$Area
group.4a<-as.character(group.4a)
group.4a<-as.numeric(group.4a)
g4_TA<-sum(group.4a)
#
####VE index calculation####
#Pi
g1_pi<-((g1_TA/A_TA)*100)/100
g2_pi<-((g2_TA/A_TA)*100)/100
g3_pi<-((g3_TA/A_TA)*100)/100
g4_pi<-((g4_TA/A_TA)*100)/100
#
#Ln Pi
g1_lnp<-log(g1_pi)
ifelse(g1_pi>0,g1_lnp,g1_lnp<-0)
g2_lnp<-log(g2_pi)
ifelse(g2_pi>0,g2_lnp,g2_lnp<-0)
g3_lnp<-log(g3_pi)
ifelse(g3_pi>0,g3_lnp,g3_lnp<-0)
g4_lnp<-log(g4_pi)
ifelse(g4_pi>0,g4_lnp,g4_lnp<-0)
#
#"-(Pi * ln Pi)"
g1_plp<- -(g1_pi*g1_lnp)
g2_plp<- -(g2_pi*g2_lnp)
g3_plp<- -(g3_pi*g3_lnp)
g4_plp<- -(g4_pi*g4_lnp)
#

```

```

#I
Ind<-g1_plp+g2_plp+g3_plp+g4_plp
#
#VE
VE<-Ind/log2(4)
#
###output row###
IMS<-c(ID,VE)
IMS<-as.data.frame(t(IMS))
ID<-ID+1                #ID vaule plus 1
#
###write to output table###
#
results<-merge(IMS,results,all=TRUE)
#
rm(tree.subset,max,req,req_2,req_3,req_4,req_5,AA,A_TA,group.1,group.1a,group.2,group.2
a,group.3,group.3a,group.4,group.4a,g1_TA,g2_TA,g3_TA,g4_TA,g1_pi,g2_pi,g3_pi,g4_pi,
g1_lnp, g2_lnp, g3_lnp, g4_lnp, g1_plp, g2_plp, g3_plp, g4_plp, Ind, VE, IMS)
#
}

else
{
###If there is <1 values for a Cell ID###
VE<-0
###output row###
IMS<-c(ID,VE)
IMS<-as.data.frame(t(IMS))
ID<-ID+1                #ID vaule plus 1
#
###write to output table###
#
results<-merge(IMS,results,all=TRUE)
#
rm(tree.subset,VE,IMS)
#
}
#####
}#end loop
colnames(results)<-c("ID", "VE")
write.csv(results,file="Tiffs_VE_index.csv", row.names=T)
write.dbf(results, file="Tiffs_VE_index2.dbf")

```

C.11 Aggregation index calculated from ITCR script– R code developed to estimate the Clark-Evans Aggregation Index (Clark and Evans, 1954) and test statistics (Vorcak et al., 2006) from TIFFS point data.

Overview: Custom R code was developed to estimate the Clark-Evans Aggregation Index (R) (Clark and Evans, 1954) and the associated test statistic (t) (Vorcak et al., 2006). The following R code follows the necessary steps of calculating the index based on the spatial arrangement of trees (TIFFs LiDAR derived Tree crown centroids). The code describes how the distances were calculated in between each of the trees encountered in one of the 30x30m grid cells, identified by ID number. The distance to its nearest neighbour was then extracted, an average of which was calculated for the cell area. Additionally, a total of the trees within the plot were calculated. The formulae to calculate the R index and t test statistic were then performed using the aforementioned inputs. An additional clause was added to account for grid cells where no trees were located.

```
###Library foreign###
library(foreign)#Allows the import of .dbf files
###load data###
tree.data<- read.dbf("Tiffs_pnts_XY.dbf") #Tiffs derived tree crown centroid locations
tree.data<-tree.data[c(11:13)]
#create results table#
results<- array(NA,c(0,4))#results table has 4 columns
#
###Distance function###
f<-function(w) apply(xy,1,function(v) sum((v-w)^2))
###select by polygon ID###
ID<-3033
###Approximate expected spacing###
SP<-2 #expected distance between trees (m)
while (ID <= 44030) {#begin to loop through cell ID no.
  print(ID)
  tree.subset<-tree.data[which(tree.data$CELLID==ID),]
  if (nrow(tree.subset)>1) {#clause for if no trees are encountered
    #
    x<-tree.subset$POINT_X
    y<-tree.subset$POINT_Y
    #
    ###compute all distances between distinct points###
    #
    xy<-cbind(x,y)
    distances<-apply(xy,1,f)
    distances<-sqrt(distances)
    #minimum distancebetween neighbour trees#
    minOfColumns=apply(distances, 2, function(x) min(x[x!=0]) )
    #
    #count trees#
    POINT_X<-tree.subset$POINT_X
    POINT_X<-as.data.frame(table(POINT_X))
    no.tree<-sum(POINT_X$Freq)
    #
    ###Aggregation Index (R)###
    R<-((1/no.tree)*sum(minOfColumns))/(0.5*sqrt(((30*30)/no.tree)))
  }
```

```

#
###Stochastic Test (t)###
t<-((mean(minOfColumns))-SP)/0.26136/sqrt((no.tree*(no.tree/(30*30)))
#
#Return the distance statistics#
IMS<-c(ID,no.tree,R,t) #output row
IMS<-as.data.frame(t(IMS))
ID<-ID+1 #ID value plus 1
#
###write to output table###
#
results<-merge(IMS,results,all=TRUE)
#
rm(x,y,xy,distances,POINT_X,no.tree,r,t,IMS)
}
else#clause for if no trees are encountered
{
#If no trees, zeros input for index values#
R<-0
t<-0
no.tree<-0
###output row###
IMS<-c(ID,no.tree,R,t) #output row
IMS<-as.data.frame(t(IMS))
ID<-ID+1 #ID value plus 1
#
###write to output table###
#
results<-merge(IMS,results,all=TRUE)
#
rm(x,y,xy,distances,POINT_X,no.tree,r,t,IMS)
#
}
#####
}#end loop
colnames(results)<-c("ID", "no_tree", "R", "t")#updates column names
write.csv(results,file="TIFFS_aggregation.csv", row.names=T)
write.dbf(results, file="TIFFS_aggregation2.dbf")

```

Appendix D – R statistics script listing

D.1 Predictor variable reduction R script

Overview: The AICc procedure (using the MuMin package – available from: <http://cran.r-project.org/web/packages/MuMin/index.html>) is used with a subset of 6 random predictor variables in each of the 200,000 iterations. The results of which were a table of how many times each of the predictors was considered statistically significant.

```
#Load library#
library (MuMin)           #For stepwise AIC-c#
#
#Input Dataset - column 1 is ID measures#
Resp<- read.csv("Feild_data.csv")      #Responce variables#
Pred<- read.csv("LiDAR_DR_data.csv")    #Predictor variables (LiDAR - ht. and int.)#
n<-names(Pred)                        #Records column names from 'Pred'#
#
#Set up simulation settings
nsims<- 300000                     #No. of loops#
nv<- 196                           #No. of total predictor variables DR LiDAR#
nused<- 6                           #No. of variables in each sample#
i<-1                                #counter is set to 1 (begin)#
results<- array(NA,c(0,nv))        #output results table (Rows:0 & Columns:95; all cells with
value 'NA')#
colnames(results)<-c(n)             #assigns same column names as Lidar.CSV file#
results<-as.matrix(results)#convert data frame to matrix
#
a <- subset(Resp, select=c("Feild_var")) #subset Feildwork table (i.e. 1 Metric only)#
#
while (i<= nsims) {                 #***start of loop 1***(for loop runs out of memery for large#
vector)#
#
b<-sample(Pred, nused)              #Random samples (6)#
d<-cbind(a,b)                       #merges Feild and RS subsets#
#
print(i)#display iteration no.#
####Regression####
fit<-lm(Field_var~., data=d)         #initial linear model(update dependent on
field# metirc)#
model<-dredge(fit)                   #AIC-c#
s<-(get.models(model, 1))[[1]]       #extract best model#
sa<-anova(s)                         #performsanova test on linear model#
sb<-sa[5]                            #pulls out Pr(>F) values i.e. significance#
x<-as.atrrix(t(sb))                  #Flips rows and columns, and converts to
matrix
x[is.na(x)]<-999                     #recodes NA to 999#
#
#If p values are <0.05 code to 1, else set to 0. (Will generate a warning, yet results are ok)#
for (j in 1:(ncol(x))) {if (x[j]<0.05) {x[j]<-1} else {x[j]<-0}}
#
results<-merge(x,results,all=TRUE)   #adds output of loop to results table
#
```



```
rm(b,d,fit,model,s,sa,sb,j,x,iter,final)           #Removes data within loop#  
i<-i+1                                             #Add 1 to the counter#  
} #End loop 1#  
#print(results)                                   #Prints a results table (not nessecary given size)#  
res_counts<-colSums(results, na.rm=TRUE, dims=1)   #will produce a total for each  
column#  
write.csv(res_counts,file="output.csv", row.names=T) #outputs result#
```

D.2 AICc implementation using reduced dataset R script

Overview: AICc was applied to the reduced number of predictor variables. This applies the stepwise AICc procedure, and selects the ‘best model’ (i.e. the highest alpha value), and performs various diagnostic tests. The three packages required are:

1. MuMin (<http://cran.r-project.org/web/packages/MuMIn/index.html>);
2. faraway (<http://cran.r-project.org/web/packages/faraway/>);
3. and perturb (<http://cran.r-project.org/web/packages/perturb/>).

```
#Load Library
library(MuMin)    #for stepwise AICc
library(faraway)  #for calculating VIF
library(perturb)  #for condition index
#
#Input Dataset
Resp<- read.csv("Feild_all.csv")    #Response variables
Pred<- read.csv("ave_space_dr_data.csv") #Predictor variables (LIDAR - ht. and int.)
#
a <- subset(Resp, select=c("Mean_dbh")) #subset Fieldwork table
b <- subset(Pred, select=c("TIF_m_area", "TIF_TreeNo", "TIF_space", "hb085i_2",
"hb075i_2", "hb090i", "SIN_Kur", "SIN_Ske", "SIN_MAD", "hb070i_2", "p075_2",
"SIN_MAD_2", "p070_2")) #List of sample attributes as identified in part 1
#
d<-cbind(a,b) #merges Field and RS subsets
####Regression###
fit<-lm(Mean_dbh~., data=d) #initial linear model(update dependent onfieldmetric)
model<-dredge(fit, trace="TRUE",m.max=6) #AIC-c with a max of 6 variables in the
#subset(model, delta<4) #displays models of delta value less than 4
s<-(get.models(model, 1))[[1]] #extract ‘best’ model (lowest AICc weight)
#
#Tests#
ss<-summary(s) #summary - R-squ. and t-test
#
sa<-anova(s) #performesanova test (f-test)
#
ss #display summary
sa #display anova
#
layout(matrix(1:4,2,2))
plot(s) #Displays diagnostic graphs
#
vif(s) #computes VIF values (<10 are ok, =<3 better)
#
colldiag(s)#computes condition index (<15 best)
#
#outputs#
#
sc<-coef(s) #extract coefficients of model
#
write.csv(sc,file="Mean_dbh_DR_AICc_coef001.csv") #output intercepts and coefs
```

```

write.csv(sa,file="Mean_dbh_DR_AICc_anova001.csv", row.names=T)    #outputs anova
table
capture.output(ss,file="Mean_dbh_DR_AICc_sum001.txt")    #outputs values of summary
#
lemon<-subset(model, delta<4)    #extracts all models of delta value less than 4
capture.output(lemon,file="Mean_dbh_DR_AICc_weights001.txt")    #outputs values of
AICc models
#
#output graphs#
jpeg("D:/R_DR/graphs/Mean_dbh_DR_001.jpg")
layout(matrix(1:4,2,2))
plot(s)
dev.off()
#

```

D.3 R script for combining datasets– AICc implementation using reduced dataset R code for combining Hyperspectral and LiDAR datasets (a modification of code listed in C2).

Overview: AICc was applied to the reduced number of predictor variables for both LiDAR and Hyperspectral metrics. This applies the stepwise AICc procedure, and selects the ‘best model’ (i.e. the highest alpha value). And performs various diagnostic tests.

```
#Load Library
library (MuMin)    #for stepwise AICc...
library (faraway)  #for calculating VIF...
library (perturb)  #for condition index...
#
#Input Dataset
Resp<- read.csv("Feild_all.csv")          #Response variables
Pred.dr<- read.csv("DR_lidar_all_int.csv") #Predictor variables for LiDAR
Pred.h<- read.csv("hyper_met.csv")        #Predictor variables for Hyperspectral
#
a <- subset(Resp, select=c("Mean_dbh"))    #subset Fieldwork table
#
####13 max each for pred subsets####
b <- subset(Pred.dr, select=c("TIF_m_area", "TIF_TreeNo", "hb085i_2", "hb075i_2",
"hb090i", "SIN_Kur", "SIN_Ske", "SIN_MAD", "hb070i_2", "p075_2"))
      #List of sample attributes as identified in part 1
#
c <- subset(Pred.h, select=c("rendvi1_STD", "sri2_STD", "mrendvi1_STD", "no_tree",
"mresri2_STD", "rendvi2_STD", "wbi1_Range", "sipi1_Mean", "cri1_Mean"))
      #List of sample attributes as identified in part 1
#
d<-cbind(a,b,c)          #merges Field and RS subsets
#
####Regression####
fit<-lm(Mean_dbh~., data=d)    #initial linear model(update dependent onfield metric)
model<-dredge(fit, trace="TRUE",m.max=6) #AIC-c with a max of 6 variables in the
#subset(model, delta<4)        #displays models of delta value less than 4
s<-(get.models(model, 1))[[1]] #extract ‘best’ model (lowest AICc weight)
#
#Tests#
ss<-summary(s)              #summary - R-squ. and t-test
#
sa<-anova(s)                #performesanova test (f-test)
#
ss #display summary
sa #display anova
#
layout(matrix(1:4,2,2))
plot(s)                     #Displays diagnostic graphs
#
vif(s) #computes VIF values (<10 are ok, =<3 better)
#
colldiag(s)#computes condition index (<15 best)
#
```

```

#outputs#
#
sc<-coef(s)          #extract coefficients of model
#
write.csv(sc,file="Mean_dbh_HDR_AICc_coef001.csv")    #output intercepts and coefs
write.csv(sa,file="Mean_dbh_HDR_AICc_anova001.csv", row.names=T)    #outputs anova
table
capture.output(ss,file="Mean_dbh_HDR_AICc_sum001.txt")    #outputs values of summary
#
lemon<-subset(model, delta<4)    #extracts all models of delta value less than 4
capture.output(lemon,file="Mean_dbh_HDR_AICc_weights001.txt")    #outputs values of
AICc models
#
#output graphs#
jpeg("D:/R_DR/graphs/Mean_dbh_HDR_001.jpg")
layout(matrix(1:4,2,2))
plot(s)
dev.off()
#

```

Appendix E – Estimation of field plot level metrics and index construction

E.1 – R script example of estimation of field plot level metrics from Remote Sensing data

Overview: once the relevant remote sensing metrics have been extracted for the spatial extents of the 2012 validation fieldwork polygons the regression model equations can be applied in order to predict field plot level information. The following demonstrates the calculation of 10 metric using a combined FW LiDAR and Hyperspectral dataset. The script uses a loop to calculate the 10 metrics for each polygon individually for 20 iterations (i.e. 20 plots). This code is part 1 of 4.

```
#####
#Field level metrics estimations from linear multiple regression (AICc)#
#####
#
###load data###
Input.1 <- read.csv("fw_metrics.csv") #FW LiDAR metrics
Input.2 <- read.csv("H_metrics.csv") #Hyper. Metrics
tree.data<- cbind(Input.1, Input.2)#simple merge as tables already organised
#create results table#
results<- array(NA,c(0,11))
#
###select by field polygon ID###
ID<-22
while (ID <= 41) {
  print(ID)
  tree.subset<-tree.data[which(tree.data$Site_no==ID),]
  #
  #####Regression models#####
  #
  ###[1] no. trees = kurht_2 + sri2_Min + TIF_m_area ###
  b_0 <- 42.0392016716989
  x_1 <- tree.subset$kurht_2
  b_1 <- 5.56287416809215
  x_2 <- tree.subset$sri2_Min
  b_2 <- 0.933545350587574
  x_3 <- tree.subset$TIF_m_area
  b_3 <- -0.370085054901359
  n_tree<- b_0+b_1*x_1+b_2*x_2+b_3*x_3
  rm(b_0, x_1, x_2, b_1, b_2, x_3, b_3)
  n_tree<- round(n_tree)
  #
  ###[2] no. Native trees = sri2_Min + wd_p45###
  b_0 <- 47.9239254629305
  x_1 <- tree.subset$arvi2_Min
  b_1 <- 1.450808158186
  x_2 <- tree.subset$wd_p45
  b_2 <- -1.24794352586117
  n_nat_tree<- b_0+b_1*x_1+b_2*x_2
  rm(b_0, x_1, x_2, b_1, b_2)
  n_nat_tree<- round(n_nat_tree)
  if (n_nat_tree>n_tree ) {n_nat_tree<- n_tree }
  #
  ###[3] Perc. Native trees = arvi1_Min + sri1_Sum + wbi2_STD###
```

```

b_0 <- 145.677977047386
x_1 <- tree.subset$arvil_Min
b_1 <- -231.675235956868
x_2 <- tree.subset$sri1_Sum
b_2 <- -0.0045824390091343
x_3 <- tree.subset$wbi2_ST
b_3 <- -568.864097646821
P_nat_tree<- b_0+b_1*x_1+b_2*x_2+b_3*x_3
rm(b_0, x_1, b_1, x_2, b_2, x_3, b_3)
if (P_nat_tree< 0) {P_nat_tree<- 0}
if (P_nat_tree> 100) {P_nat_tree<-100}
#
####[4] Tree_spec = rendvil_Max + rendvil_STD + sipi1_Range###
b_0 <- 6.04046099373652
x_1 <- tree.subset$rendvil_Max
b_1 <- -9.43249299222572
x_2 <- tree.subset$rendvil_STD
b_2 <- 55.420126438021
x_3 <- tree.subset$sipi1_Range
b_3 <- -1.28912967751252
tree_spec<- b_0+b_1*x_1+b_2*x_2+b_3*x_3
rm(b_0, x_1, x_2, b_1, b_2, x_3, b_3)
tree_spec<- round(tree_spec)
#
####[5] Avg_space = pri2_STD + TIF_Space + varht_2###
b_0 <- 2.46610685169097
x_1 <- tree.subset$pri2_STD
b_1 <- -167.042177706407
x_2 <- tree.subset$TIF_Space
b_2 <- 0.462374249625309
x_3 <- tree.subset$varht_2
b_3 <- 0.024540558661383
Avg_space<- b_0+b_1*x_1+b_2*x_2+b_3*x_3
rm(b_0, x_1, x_2, b_1, b_2, x_3, b_3)
#
####[6] STD_space = sri2_Range + TIF_m_area###
b_0 <- 0.553684220907678
x_1 <- tree.subset$sri2_Range
b_1 <- 0.0146595522853885
x_2 <- tree.subset$TIF_m_area
b_2 <- 0.0193965203557637
STD_space<- b_0+b_1*x_1+b_2*x_2
rm(b_0, x_1, x_2, b_1, b_2)
#
####[7] Mean_DBH =TIF_m_area + varngwid_2###
b_0 <- 47.6793961599107
x_1 <- tree.subset$TIF_m_area
b_1 <- 0.339219056953958
x_2 <- tree.subset$varngwid_2
b_2 <- -0.0141330321444415
Mean_DBH<- b_0+b_1*x_1+b_2*x_2
rm(b_0, x_1, x_2, b_1, b_2)
#
####[8] StDev_DBH = amp_p65 + TIF_m_area + wd_p25_2###
b_0 <- 14.1482758025794

```



```

x_1 <- tree.subset$amp_p65
b_1 <- -1.05669943050351
x_2 <- tree.subset$TIF_m_area
b_2 <- 0.194681260577237
x_3 <- tree.subset$wd_p25_2
b_3 <- 0.959834961240511
StDev_DBH<- b_0+b_1*x_1+b_2*x_2+b_3*x_3
rm(b_0, x_1, x_2, x_3, b_1, b_2, b_3)
#
###[9] Total_BA = meanht + rendvil_Max + TIF_Space ###
b_0 <- 5.96011685133491
x_1 <- tree.subset$meanht
b_1 <- 0.207343475415511
x_2 <- tree.subset$rendvil_Max
b_2 <- -5.69841431501339
x_3 <- tree.subset$TIF_Space
b_3 <- -0.202867422
Total_BA<- b_0+b_1*x_1+b_2*x_2+b_3*x_3
rm(b_0, x_1, x_2, b_1, b_2, x_3, b_3)
#
###[10] B_tree = kurgamp_2 + mednght + TIF_Space###
b_0 <- -40.7244581979019
x_1 <- tree.subset$kurgamp_2
b_1 <- -6.25287592951413
x_2 <- tree.subset$mednght
b_2 <- 4.03570454646313
x_3 <- tree.subset$TIF_Space
b_3 <- 3.72881553821154
B_tree<- b_0+b_1*x_1+b_2*x_2+b_3*x_3
rm(b_0, x_1, x_2, b_1, b_2, x_3, b_3)
if (B_tree< 0) {B_tree<- 0}
if (B_tree> 100) {B_tree<-100}
#
###create output row for results table###
IMS <- c(ID, n_tree, n_nat_tree, P_nat_tree, tree_spec, Avg_space, STD_space, Mean_DBH,
StDev_DBH, Total_BA, B_tree)
IMS<-as.data.frame(t(IMS))
#merges to results table#
results<-merge(IMS,results,all=TRUE)
#
ID<-ID+1
#####
}
#assigns column names to results table#
colnames(results)<-c("ID", "no_tree", "no_nat_tree", "perc_nat_tree", "tree_spec", "Avg_space",
"STD_space", "Mean_DBH", "StDev_DBH", "Total_BA", "B_tree")
write.csv(results,file="Val_part1_10_f2012.csv", row.names=T)

```

E.2 – R script example of estimation of condition indices from derived field plot level metrics

Overview: Once the metrics have been calculated (and the best estimates selected) the conditional indices can be calculated. Again the calculations are applied in a loop, where each loop-iteration extracts metrics for one field site and calculates the four condition indices. These indices (in order of appearance) are: Score index 1; Complexity Index, Diameter Differentiation index and Complex Stand diversity index.

```
####load data###
metric.data <- read.csv("Best_DR_FW_H.csv")
#additional metrics#
Input.A <- read.csv("fw_metrics.csv")
Input.B <- read.csv("Fld_lvl_Class_metrics.csv")
Input.C <- read.csv("TIFFS_CSDI_part_F2012.csv")
#
#create results table#
results <- array(NA,c(0,5))
#
####select by polygon ID###
ID<-22
while (ID <= 41) {
  print(ID)
  metric.subset<-metric.data[which(metric.data$ID==ID),]
  #
  ####Score index 1###
  #
  ####Indicators 1-17###
  #[1] No. of trees (ha-1) – convert to count per hectare#
  no.trees <- round((metric.subset$no_tree/9)*100)
  #[2] Shannon index#
  SH_count <- metric.subset$SH_count
  #[3] Basal Area (m2 ha-1)#
  Total_BA <- (metric.subset$Total_BA/9)*100
  #[4] Mean DBH (cm)#
  Mean_DBH <- metric.subset$Mean_DBH
  #[5] STD of DBH (cm)#
  StDev_DBH <- metric.subset$StDev_DBH
  #[6] Percentage of big trees (>40cm DBH)#
  big_tree <- metric.subset$big_tree
  if (big_tree < 0) {big_tree <- 0}
  if (big_tree >100) {big_tree <- 100}
  #[7] Mean Height#
  FW.subset<-Input.A[which(Input.A$Site_no==ID),]
  mean_ht <- FW.subset$TIF_HT
  #
  #[8] No. of total saplings (ha-1)#
  No_saplings <- round((round(metric.subset$No_saplings)/9)*100)
  #[9] No. of native saplings (ha-1)#
  No_Nat_sap <- round((round(metric.subset$No_Nat_sap)/9)*100)
  if (No_saplings < No_Nat_sap) {No_Nat_sap <- No_saplings}
  #check result of if...#
  #
  #[10] Vol. downed dead wood (m3 ha-1)#
  D_ddwvol <- metric.subset$D_ddwvol*100
```

```

#[11] Downed dead wood decay class#
D_DDW_class <- metric.subset$D_DDW_class
if (D_ddwvol <= 0) {D_DDW_class <- 0}
#
#[12] Volume of snag (m3 ha-1)#
D_snagvol <- (metric.subset$D_snagvol/9)*100
#[13] Snag Decay class#
D_Snag_class <- metric.subset$D_Snag_class
if (D_snagvol <= 0) {D_Snag_class <- 0}
#
#[14] No. of total seedlings (ha-1)#
No_seedlings <- round(metric.subset$No_seedlings*100)
#[15] No. of native seedlings (ha-1)#
No_nat_seed <- round(metric.subset$No_nat_seed*100)
if (No_saplings < No_Nat_sap) {No_seedlings <- No_nat_seed}
#[16] Shannon index for native seedlings#
Nat_seedlings_SH <- metric.subset$Nat_seedlings_SH
if (Nat_seedlings_SH < 0) {Nat_seedlings_SH <- 0}
#[17] No. of ground veg. species#
No_vascular <- round(metric.subset$No_vascular)
#
###Index (score no.1) construction###
#[1]
if (no.trees >= 222) {I.1 <- 1} else {I.1 <- 0}
#[2]
if (SH_count >= 0.87) {I.2 <- 1} else {I.2 <- 0}
#[3]
if (Total_BA >= 23) {I.3 <- 1} else {I.3 <- 0}
#[4]
if (Mean_DBH >= 32) {I.4 <- 1} else {I.4 <- 0}
#[5]
if (StDev_DBH >= 14) {I.5 <- 1} else {I.5 <- 0}
#[6]
if (big_tree >= 7) {I.6 <- 1} else {I.6 <- 0}
#[7]
if (mean_ht >= 17) {I.7 <- 1} else {I.7 <- 0}
#[8]
if (No_saplings >= 91) {I.8 <- 1} else {I.8 <- 0}
#[9]
if (No_Nat_sap >= 91) {I.9 <- 1} else {I.9 <- 0}
#[10]
if (D_ddwvol >= 26) {I.10 <- 1} else {I.10 <- 0}
#[11]
if (D_DDW_class >= 0.5) {I.11 <- 1} else {I.11 <- 0}
#[12]
if (D_snagvol >= 16) {I.12 <- 1} else {I.12 <- 0}
#[13]
if (D_Snag_class >= 0.4) {I.13 <- 1} else {I.13 <- 0}
#[14]
if (No_seedlings >= 63219) {I.14 <- 1} else {I.14 <- 0}
#[15]
if (No_nat_seed >= 63219) {I.15 <- 1} else {I.15 <- 0}
#[16]
if (Nat_seedlings_SH >= 0.89) {I.16 <- 1} else {I.16 <- 0}
#[17]

```

```

if (No_vascular >= 0.89) {I.17 <- 1} else {I.17 <- 0}
#Sum index values#
score_index_1 <-
I.1+I.2+I.3+I.4+I.5+I.6+I.7+I.8+I.9+I.10+I.11+I.12+I.13+I.14+I.15+I.16+I.17
#####
#
### Complexity Index (HC) ###
tree_spec <- round(metric.subset$tree_spec)
HC_index <- (mean_ht * Total_BA * no.trees * tree_spec)/10000
#####
#
### Diameter Differentiation index###
DDI_index <- 1-(1/no.trees)*(metric.subset$Sum_Diam_Diff)
#
#####
#
### Complex Stand diversity index ###
#[1] tree species composition
ecog.subset<-Input.B[which(Input.B$ID==ID),]
##
CSDI.1 <- log(tree_spec)*(1.5-
(((ecog.subset$max_pop/no.trees)*100)/100)+(((ecog.subset$min_pop/no.trees)*100)/100))
#[3]Index of spatial distribution – with correction#
STD_space <- metric.subset$STD_space
CSDI.3.fb <- 1/((((CSDI.subset$Min.d1-(2*STD_space)) + (CSDI.subset$Min.d2-
(2*STD_space)) + (CSDI.subset$Min.d3)-(2*STD_space))/3)+(((CSDI.subset$Max.d1-
(2*STD_space)) + (CSDI.subset$Max.d2-(2*STD_space)) + (CSDI.subset$Max.d3)-
(2*STD_space))/3))
#
CSDI.3.vb <- (1-((CSDI.subset$Min.d1-(2*STD_space)) + (CSDI.subset$Min.d2-
(2*STD_space)) + (CSDI.subset$Min.d3-(2*STD_space)))/((CSDI.subset$Max.d1-(2*STD_space))
+ (CSDI.subset$Max.d2-(2*STD_space)) + (CSDI.subset$Max.d3-(2*STD_space))))*CSDI.3.fb
#
#[4]Index of crown differentiation
Avg_base_ht <- metric.subset$Avg_base_ht
StDev_base_ht<- metric.subset$StDev_base_ht
if (Avg_base_ht < 2) {StDev_base_ht <- 0}
if (Avg_base_ht < StDev_base_ht) {StDev_base_ht <- 0}
CSDI.4 <- (1-log(Avg_base_ht-StDev_base_ht)/4)+((CSDI.subset$Max.cr1+
CSDI.subset$Max.cr2)/(CSDI.subset$Min.cr1 + CSDI.subset$Min.cr2))
#[5] Calculate index from components 1-4
p <- 4
q <- 3
CSDI.5c <- p * CSDI.1 + q * metric.subset$IOVS_1 + CSDI.3.vb + CSDI.4
#
IMS <- c(ID, score_index_1, HC_index, DDI_index, CSDI.5c)
IMS<-as.data.frame(t(IMS))
results<-merge(IMS,results,all=TRUE)
#
ID<-ID+1
#####
}
#Add column headings#
colnames(results)<-c("ID", "score_index_1", "HC_index", "DDI_index", "CSDI")
#output results table#

```

```
write.csv(results,file="Val_indexes_f2012.csv", row.names=T)
```

Appendix F–Comparison of mapping of forest metrics from different inputs

F.1 – A comparison of mapped remote sensing modelled forest metrics

As stated in section 4.7, the regression models could be applied for the entirety of the study site, provided there was coincident airborne remote sensing data coverage to provide input metrics. A subset of 10 out of the total of 39 field-level metrics are presented here. The examples of the ten field-level metrics within each 30x30m area illustrated are: (i) the number of tree stems; (ii) the number of tree species encountered; (iii) SH diversity index; (iv) the total basal area; (v) the mean DBH; (vi) standing deadwood volume; (vii) standing deadwood decay class; (viii) the number of saplings; (ix) the number of seedlings; and (x) the number of seedling species.

The mapping subset will centre upon the Frame Wood area, with elements of Tantany Woods, Frame-Heath and Hawkhill inclosures visible. This area contains a mixture of semi-ancient woodland, such as those located in Frame and Tantany Woods, and managed plantation woodland, such as that in the Frame-Heath and Hawkhill inclosures. Figure F.1 illustrates the subset area comprising a 1x1m resolution CHM, and the primary species as identified by FC inventory data.

Those Hyperspectral, DR LiDAR or composite DR and hyperspectral regression models reported in Chapter 8.4 were used to calculate the map values. Figure F.2 presents the results of the model calculations applied over a wide area for the sample metrics i-v. The subset area for the number of tree stems (Figure F.2a), indicates higher stem densities within the plantation inclosures. A relatively even tree stem density can be observed through much of the semi-ancient woodland. The number of tree species (Figure F.2b) varies across plantation woodland compartments from low to very high. Generally, this corresponds well to the data provided by the Forestry Commission (see Figure F.1c). The SH index of diversity (see Figure F.2c) generally indicates higher diversity within the semi-ancient/deciduous woodlands. Total basal area for plantation, typically coniferous, woodland is almost double that of the semi-ancient deciduous woodland (see Figure F.2d), however the values presented are low. Patterns of values for mean DBH (see Figure F.2e) are opposite to that of total basal area, higher values are located within the semi-ancient and deciduous woodland.

Figure F.3 presents the mapping for results of a number of understorey metrics, i.e. vi-x as above. Higher standing deadwood volume values are located primarily within the semi-ancient and deciduous woodlands, whereas lower values, including zero, were found in the managed plantation areas (see Figure F.3a). This general pattern is again found concerning the standing deadwood decay class, however some deciduous areas show lower decay values. (see Figure F.3b). Higher values for the number of saplings were located within Tantany Wood and deciduous compartments located within Hawkhill (see Figure F.3c), whilst plantation coniferous area and elements of Frame Wood had lower values, this was commonly zero. The number of seedlings typically follows the same pattern (see Figure F.4d). The final sample metric, the number of seedling species, is generally higher in the semi-ancient and deciduous woodland, although there are coniferous compartments within Frame-heath and Hawkhill which exhibit higher values (see Figure F.4e).

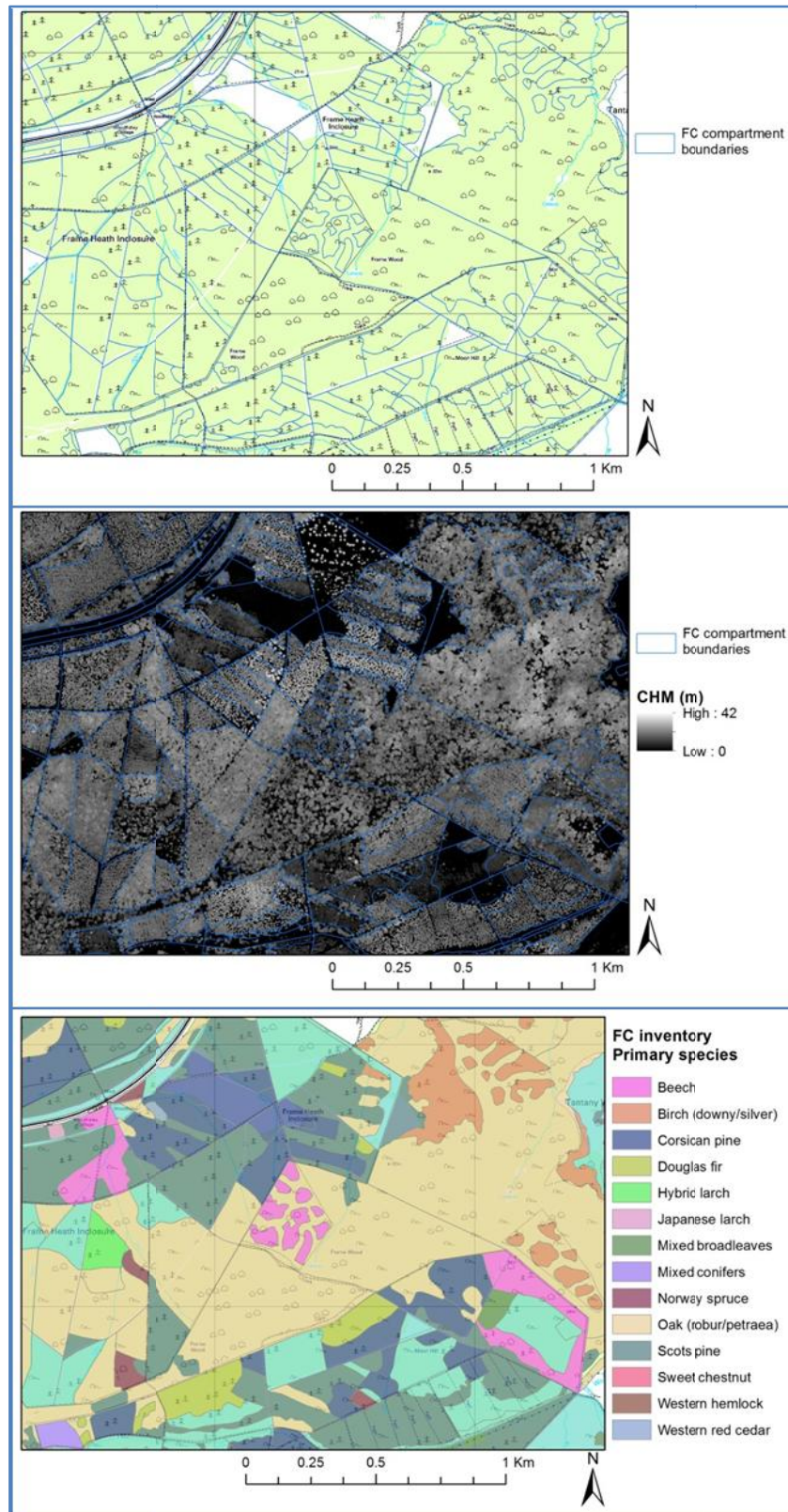


Figure F.1 – The subset area for presenting predicted field-level metrics over wide areas. Frame and Tantany Woods in addition to Frame-Heath and Hawkhillinclosures intersect this area. Ordnance survey mapping, 1x1m raster CHM, and FC inventory primary species classification are presented for reference.

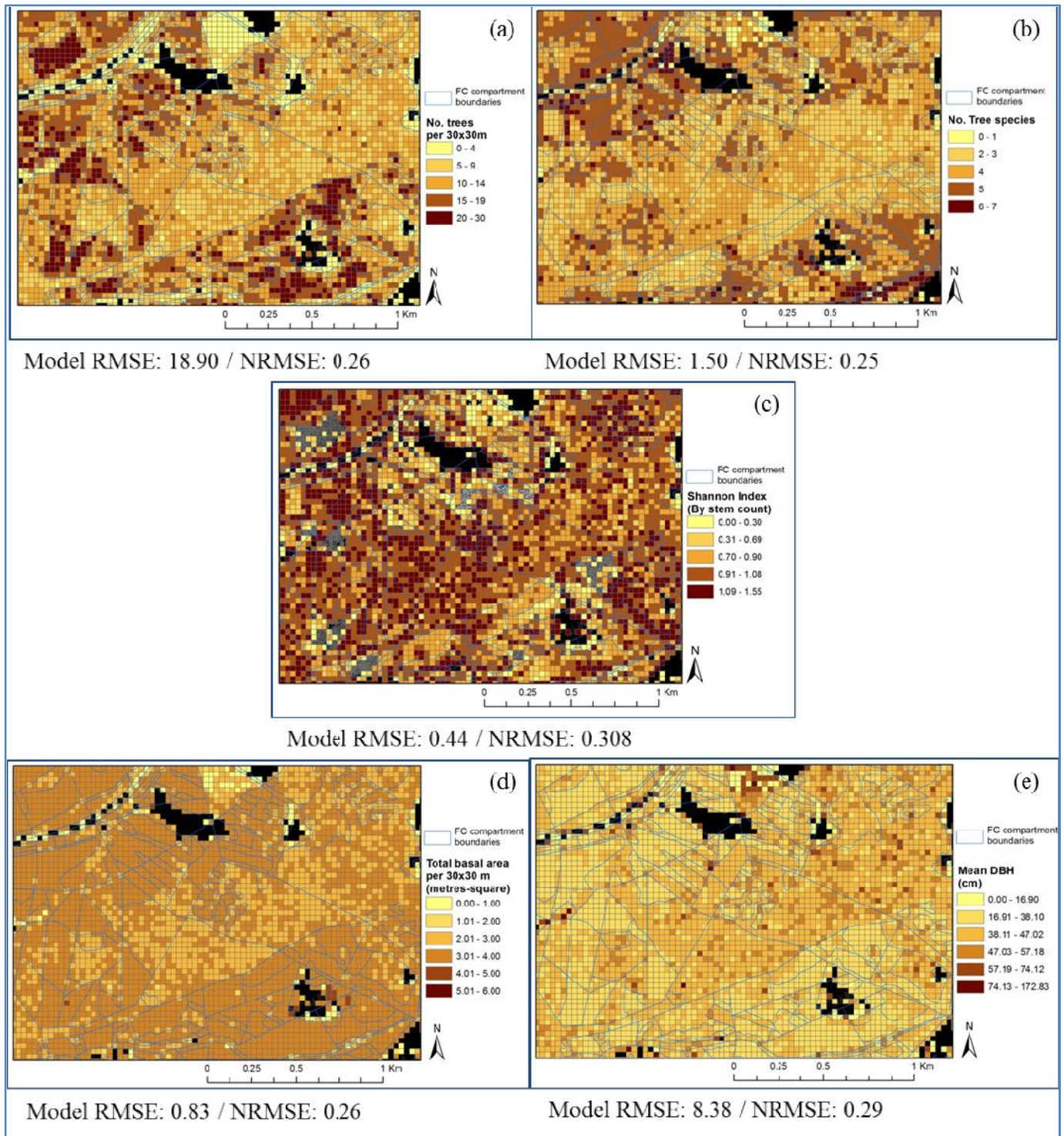


Figure F.2 – Mapped results for regression models produced by DR and hyperspectral datasets. These metrics are: (a) the number of tree stems; (b) the number of tree species encountered; (c) SH diversity index; (d) the total basal area; and (e) the mean DBH.

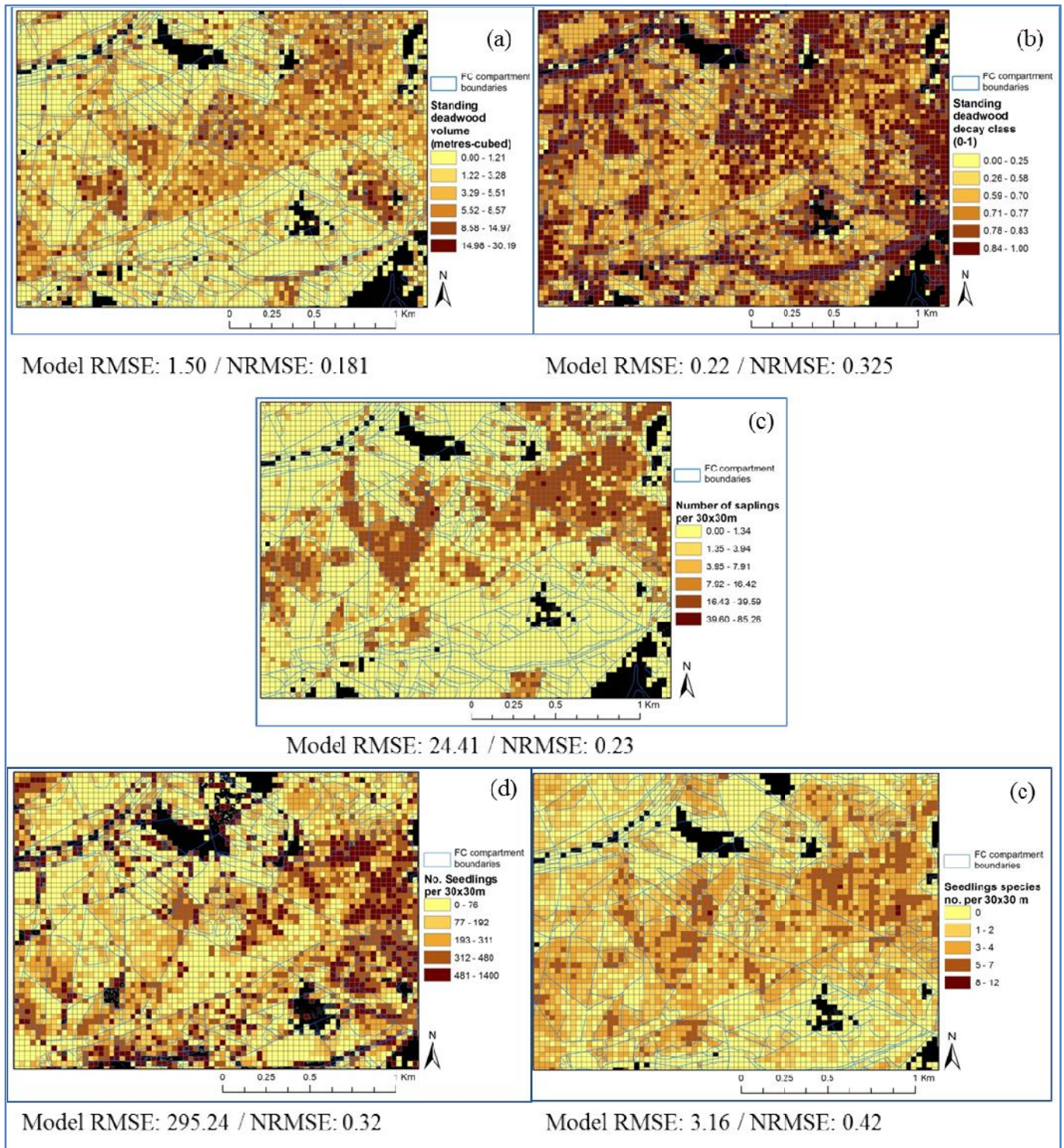


Figure F.3 - Mapped results for regression models produced by DR and hyperspectral datasets. These metrics are: (a) standing deadwood volume; (b) standing deadwood decay class; (c) the number of saplings; (d) the number of seedlings; and (e) the number of seedling species.

The hyperspectral, FW LiDAR or composite FW and hyperspectral regression models reported in chapter 8.4 were used to calculate map values. Figure F.4 presents the results of the model calculations applied over a wide area for the sample field-level metrics i-v. It should be noted that there is a line of grid cells from north to south on the eastern side of this subset area. When considering the first metric for the subset area, the number of tree stems per grid cell (see Figure F.4a), a similar pattern is observed in Figure F.2a, with higher counts present in plantation and coniferous compartments. However, higher counts are observed here. The second metric, the number of tree species (see Figure F.4b), indicates a slightly higher proportion of tree species present in semi-ancient and deciduous areas. There would seem to be fewer species detected using this dataset than that employed in Figure F.2b. The SH index of diversity also shows a very similar pattern to than observed before, however there is a greater distinction between plantation coniferous woodland and other types (see Figure F.4c). The estimates of total basal area show a very different pattern to the previous (Figure F.2d) with a different range of values. Figure F.4d has higher values in much of the semi-ancient deciduous woodland. The mean DBH has a very similar pattern to that in Figure F.2e, however higher and more variable values are detected in semi-ancient deciduous areas.

Figure F.5 presents the estimated understorey metric values for the composite hyperspectral and FW dataset. Standing deadwood volume (See Figure F.5a) shows an almost identical pattern to that depicted in Figure F.3, albeit with smaller values. The pattern of standing deadwood decay class, contains much more variability than its counterpart in Figure F.3. Conifer area show low decay class values, however deciduous areas contain a great deal of variability, for example if Tanatany Wood is contrasted with Frame Wood. The overall number of saplings (see Figure F.5c) detected by the hyperspectral and FW composite model is lower than those indicated in Figure F.3, where the overall pattern of higher values of saplings in deciduous areas is higher. Coniferous areas are however not completely devoid of values. The number of seedlings in Figure F.5d has lower values than those identified in Figure F.3d, however the same pattern of higher counts being located in semi-ancient and deciduous woodland and smaller values in plantation conifer areas. The final metric, the number of seedling species, is generally higher in the semi-ancient and deciduous woodland (see Figure F.5e). A different pattern is evident when compared with Figure F.3e, where the formers seedling species number is lower in Hawkhill deciduous compartments.

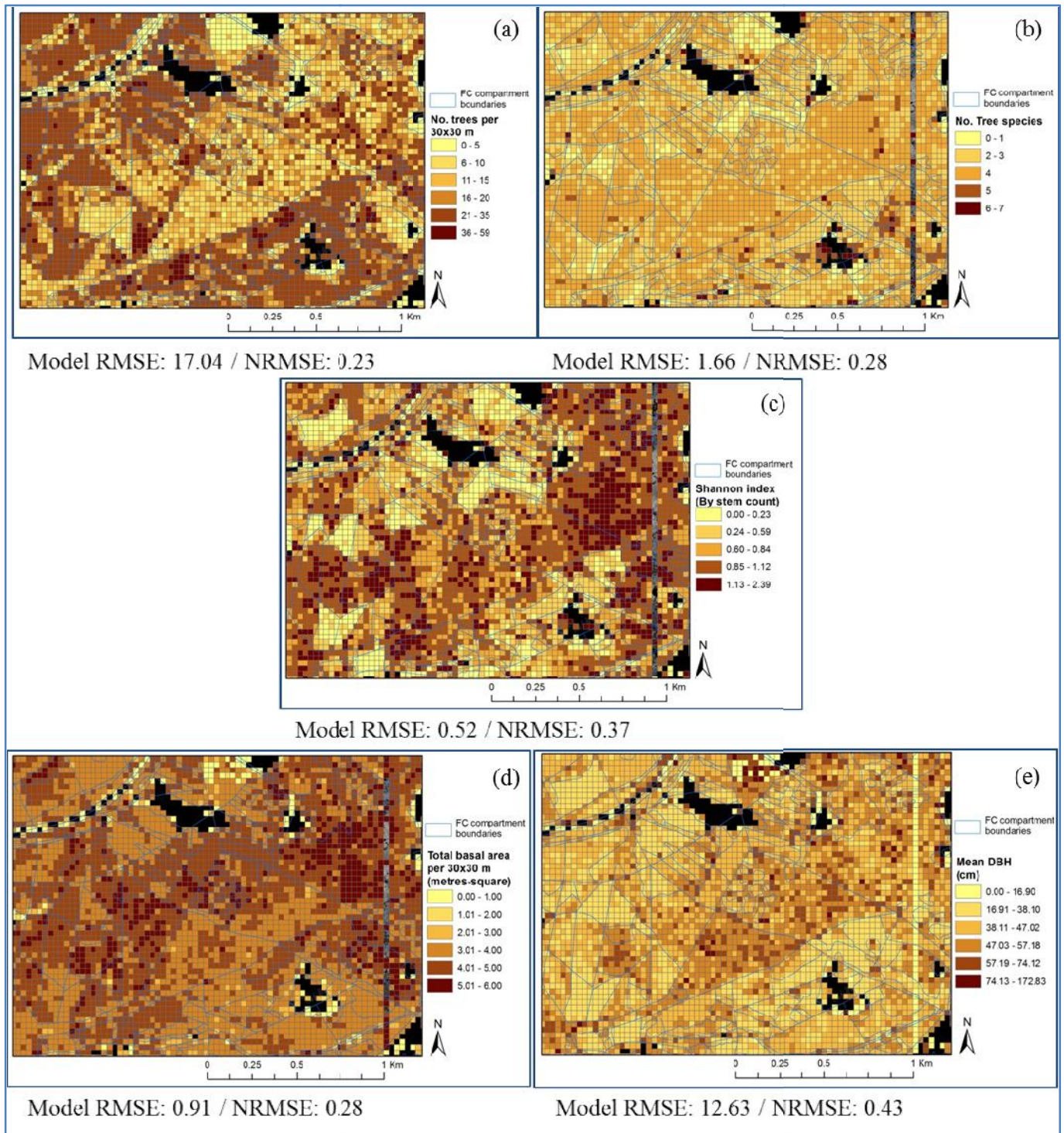


Figure F.4 – Mapped results for regression models produced by FW and hyperspectral datasets. These metrics were: (a) the number of tree stems; (b) the number of tree species encountered; (c) SH diversity index; (d) the total basal area; and (e) the mean DBH.

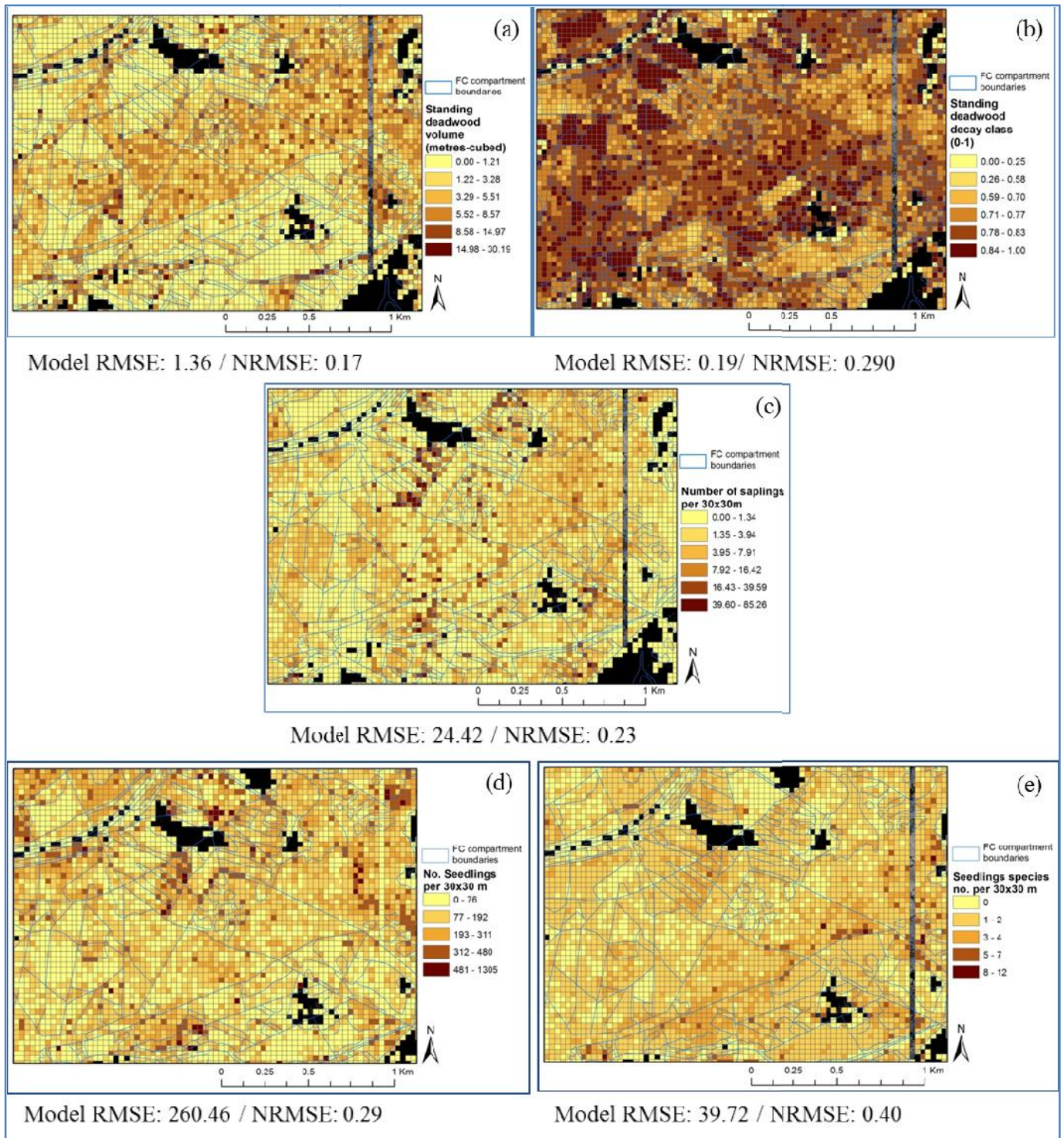


Figure F.5 - Mapped results for regression models produced by DR and hyperspectral datasets. These metrics were: (a) standing deadwood volume; (b) standing deadwood decay class; (c) the number of saplings; (d) the number of seedlings; and (e) the number of seedling species.

F.2 – A comparison of remote sensing derived condition indices

The following indices were calculated from the various remote sensing derived products relating to summarised ITC and statistically derived field plot-level metrics. The summary ITC metrics are utilised on their own for the first to indices outlined in this section. The statistically derived metrics reported here conform to three sets of model estimates, typically those identified with the smallest RMSE/NRMSE from the following: (i) hyperspectral; (ii) DR LiDAR, (iii) FW LiDAR; (iv) combined hyperspectral and DR LiDAR; and (v) combined hyperspectral and FW LiDAR dataset metrics, as identified in Chapter 8.4.

The vertical evenness (VE) index and the Clark-Evans aggregation index were both calculated from DR LiDAR ITC analysis using data generated from the TIFFS software, and as such are not mentioned in the following as no comparisons can be made.

F.2.1 The tree diameter differentiation index

The tree diameter differentiation index is a measure of the degree of variability in tree stem size within the plot. The inputs for the equation are the number of stems within the plot and an estimate of the sum of the differences between NN tree DBH values, both of which were estimated statistically. The resultant estimation of the index calculations are presented in Table F.1 for both (i) DR, hyperspectral or combined DR and hyperspectral and (ii) FW, hyperspectral or combined FW and hyperspectral datasets.

The combined DR and hyperspectral estimate of the diameter differentiation index produced an RMSE of 0.26, and NRMSE of 0.318. Plots 4 and 20, failed to produce an index value. Smaller field index values generally were best predicted. Plot 14 was underestimated by a relatively large amount. The combined FW and hyperspectral estimate of the index produced an RMSE value of 0.19 and a NRMSE of 0.23. The majority index estimations are close to the field index value. Again, plot 14 underestimated the field value, while plot 13 overestimated the field value. The estimate in Plot 12 failed.

Table F.1 – Comparison of the field derived and remote sensing derived tree diameter differentiation index

Site no.	Field calc. index	DR and/or Hyper. index	FW and/or Hyper. index
1	0.230	0.224	0.315
2	0.192	0.322	0.218
3	0.221	0.296	0.412
4	0.518	0.000	0.560
5	0.264	0.317	0.212
6	0.248	0.634	0.290
7	0.349	0.299	0.339
8	0.362	0.287	0.484
9	0.325	0.371	0.342
10	0.409	0.648	0.216
11	0.454	0.152	0.453
12	0.186	0.468	0.000
13	0.385	0.414	0.752
14	1.000	0.338	0.358
15	0.338	0.575	0.395
16	0.264	0.245	0.329
17	0.232	0.132	0.174
18	0.200	0.155	0.400
19	0.342	0.542	0.308
20	0.401	0.000	0.414
-	RMSE	0.259	0.189
-	NRMSE	0.318	0.232

F.2.2 The complexity index (HC)

The Complexity Index (HC) combines four measures of stand description: (i) number of tree species; (ii) stem number per hectare; (iii) dominant height; and (iv) the total basal area per hectare. Each of these inputs was estimated statistically through regression models, as described in the previous step. Table F.2 summarises the results of the index calculated for the field and remote sensing derived data. There is a great deal of variability in the values for all index calculations, which is reflected in the large numbers reported in the three estimates or RMSE. The metric set with the lowest overall RMSE and NRMSE is FW LiDAR and Hyperspectral, with values of 95.433 and 0.173 respectively. Strangely the index calculation computed by the metrics with the lowest RMSE, as defined in Chapter 8.4, produced an RMSE and NRMSE higher, 113.8 and 0.206 respectively. Overall values for DR LiDAR and hyperspectral models groups RMSE and NRMSE values are higher, 122.9 and 0.22 respectively. The differences in RMSE indicate the differences in the error attached to each input.

Table F.2 – Comparison of field vs. remote sensing derived HC index

Site no.	Class	Field calc. HC index	DR and/or Hyper. HC index	FW and/or Hyper. HC index	'Best' models HC index
22	<i>Deciduous</i>	45.927	19.029	27.967	29.764
23	<i>Coniferous</i>	67.462	54.079	56.309	60.028
24	<i>Deciduous</i>	88.474	71.790	163.379	83.493
25	<i>Deciduous</i>	59.631	58.924	147.677	108.126
26	<i>Deciduous</i>	64.692	11.166	33.456	22.132
27	<i>Deciduous</i>	109.911	38.423	115.222	106.815
28	<i>Coniferous</i>	109.478	29.108	65.606	42.652
29	<i>Deciduous</i>	38.563	43.739	128.338	93.241
30*	<i>Coniferous</i>	567.434	81.622	197.055	111.770
32	<i>Mixed</i>	100.279	41.201	71.523	45.996
33	<i>Coniferous</i>	137.894	77.996	166.483	75.419
34	<i>Mixed</i>	106.467	58.603	82.984	74.156
35	<i>Coniferous</i>	194.753	58.902	193.562	85.815
36	<i>Coniferous</i>	92.196	67.139	92.259	78.217
37	<i>Deciduous</i>	45.529	62.086	80.968	100.459
38	<i>Mixed</i>	120.836	51.433	111.531	81.291
39	<i>Coniferous</i>	14.841	27.737	73.994	37.516
40	<i>Deciduous</i>	31.780	66.810	58.148	71.521
41	<i>Mixed</i>	78.519	57.534	20.814	75.311
-	-	RMSE	122.911	95.433	113.842
-	-	NRMSE	0.222	0.173	0.206

F.2.3 The complex stand diversity index (CSDI)

The complex stand diversity index (CSDI) also requires a combination of inputs from a combination of direct measurements from remote sensing outputs and statistically derived outputs from regressions. It is composed of four indices relating to species composition, stem size, stem spacing, and crown dimensions. Table F.3 summarises the results of indices calculated for the field and remote sensing derived data. Plots 21 and 30 contained measurement errors in the field data and were removed as part of the comparison. Overall the RMSE and NRMSE values for all three model group estimates were very similar. The lowest RMSE value (3.281) is for DR and/or hyperspectral estimates. The index was calculated from the statistically derived estimates with the lowest RMSE, see Chapter 8.4, and has a slightly higher RMSE when the CSDI was calculated.

Table F.3 – Comparison of field vs. remote sensing derived CSDI index

Site no.	Class	Field calc. CSDI index	DR and Hyper. CSDI index	FW and Hyper. CSDI index	Best models CSDI index
22	<i>Deciduous</i>	-	8.894	11.476	11.431
23	<i>Coniferous</i>	11.499	10.322	9.888	10.609
24	<i>Deciduous</i>	12.778	10.200	13.716	10.127
25	<i>Deciduous</i>	9.920	12.503	14.724	13.168
26	<i>Deciduous</i>	10.223	8.996	12.421	10.730
27	<i>Deciduous</i>	11.465	11.460	14.292	14.045
28	<i>Coniferous</i>	6.882	9.773	11.062	11.080
29	<i>Deciduous</i>	10.461	9.941	12.558	11.110
30	<i>Coniferous</i>	-	-	-	-
31	<i>Coniferous</i>	4.273	13.531	15.607	13.077
32	<i>Mixed</i>	9.220	9.929	10.686	9.680
33	<i>Coniferous</i>	9.505	6.980	8.719	6.916
34	<i>Mixed</i>	12.390	10.452	9.883	11.237
35	<i>Coniferous</i>	10.146	8.203	12.031	8.633
36	<i>Coniferous</i>	7.417	9.342	9.447	10.259
37	<i>Deciduous</i>	8.376	10.549	11.469	12.376
38	<i>Mixed</i>	8.779	12.381	14.079	12.592
39	<i>Coniferous</i>	5.762	10.582	11.293	11.304
40	<i>Deciduous</i>	8.108	11.821	11.284	12.078
41	<i>Mixed</i>	10.640	13.860	6.737	14.620
-	-	RMSE	3.281	4.083	3.593
-	-	NRMSE	0.386	0.480	0.422

Given the nature of this combined index, the four component inputs can be broken down and analysed separately. Table F.4 summarises the differences between index inputs for the first component of the index relating to tree species composition. RMSE and NRMSE values were high, and poorly represented the index value. The index calculated from the ‘best’ estimates, i.e. those with the lowest RMSE/NRMSE identified in Chapter 8.4, performed in the middle of the other two datasets. The index incorporates estimates of tree number, species number and the relative proportions of the maximum and minimum species within each 30x30m grid cell, the latter was derived from classified ITC objects (see section 7.6). This again highlights the underestimation of tree number from the ITC delineation methods and the statistically derived number of trees.

The construction of the second index element, the index of vertical structure, was derived statistically and reported in section 8.4. The construction of the third index, the index of spatial distribution, relates to the comparison of the three largest and three smallest NN stem distances, which was reported in section 7.4.1, Table F.5 reports the comparison between field and remote sensing derived values of CSDI-3. The FW and hyperspectral derived index (which used the same inputs as those determined as the ‘best’ estimates) produces RMSE and NRMSE values lower than for using DR and hyperspectral inputs, indicating that for the latter, estimates are within 0.077 of the actual value for RMSE, and 0.453 NRMSE.

The final index input for the CSDI concerns crown differentiation, the construction of which was assessed in section 7.4.1 (for ITC derived inputs) and section 7.7 (for estimates of the lowest canopy base height). Table F.6 reports the relationship between the index values for field and remote sensing derived measures. RMSE and NRMSE values are similar for all input datasets. DR and hyperspectral dataset has the marginally lowest error, with 1.969 RMSE and 0.343 NRMSE, however the index calculated from the ‘best’ estimates, i.e. those with the lowest RMSE/NRMSE identified in Chapter 8.4, has a slightly higher error of 0.345 NRMSE (or 0.01% higher). Overall remote sensing index values typically only underestimate the field based equivalent by small amounts.

Table F.4 – Index of tree species composition comparison between field and RS derived

Site no.	Class	Field calc. CSDI-1 index	DR and Hyper. CSDI-1 index	FW and Hyper. CSDI-1 index	Best models CSDI-1 index
22	<i>Deciduous</i>	0.804	0.315	0.957	0.957
23	<i>Coniferous</i>	0.363	1.108	1.067	1.239
24	<i>Deciduous</i>	0.318	0.854	1.760	0.880
25	<i>Deciduous</i>	0.534	1.287	1.831	1.451
26	<i>Deciduous</i>	0.678	0.789	1.537	1.218
27	<i>Deciduous</i>	0.513	0.682	1.300	1.300
28	<i>Coniferous</i>	0.416	1.023	1.356	1.356
29	<i>Deciduous</i>	0.291	1.088	1.748	1.386
30	<i>Coniferous</i>	-	-	-	-
31	<i>Coniferous</i>	0.000	1.430	2.036	1.390
32	<i>Mixed</i>	0.783	0.884	1.217	0.964
33	<i>Coniferous</i>	0.452	0.487	0.936	0.468
34	<i>Mixed</i>	1.248	1.356	1.194	1.507
35	<i>Coniferous</i>	0.593	0.754	1.687	0.843
36	<i>Coniferous</i>	0.494	1.260	1.227	1.425
37	<i>Deciduous</i>	0.425	1.187	1.425	1.655
38	<i>Mixed</i>	0.490	1.286	1.791	1.419
39	<i>Coniferous</i>	0.000	1.027	1.191	1.191
40	<i>Deciduous</i>	0.463	1.390	1.256	1.458
41	<i>Mixed</i>	0.660	1.829	0.000	1.968
-	-	RMSE	0.721	1.029	0.870
-	-	NRMSE	0.577	0.825	0.697

Table F.5 – Index spatial distribution comparison between field and RS derived

Site no.	Class	Field calc. CSDI-3 index	DR and Hyper. CSDI-3 index	FW and Hyper. CSDI-3 index	Best models CSDI-3 index
22	<i>Deciduous</i>	-	-	-	-
23	<i>Coniferous</i>	0.142	0.162	0.112	0.112
24	<i>Deciduous</i>	0.134	0.066	0.057	0.057
25	<i>Deciduous</i>	0.191	0.083	0.061	0.061
26	<i>Deciduous</i>	0.217	0.049	0.062	0.062
27	<i>Deciduous</i>	0.219	0.039	0.039	0.039
28	<i>Coniferous</i>	0.092	0.073	0.084	0.084
29	<i>Deciduous</i>	0.187	0.077	0.063	0.063
30	<i>Coniferous</i>	-	-	-	-
31	<i>Coniferous</i>	0.063	0.500	0.077	0.077
32	<i>Mixed</i>	0.067	0.221	0.083	0.083
33	<i>Coniferous</i>	0.072	0.097	0.054	0.054
34	<i>Mixed</i>	0.116	0.061	0.057	0.057
35	<i>Coniferous</i>	0.088	0.053	0.055	0.055
36	<i>Coniferous</i>	0.069	0.046	0.030	0.030
37	<i>Deciduous</i>	0.084	0.066	0.053	0.053
38	<i>Mixed</i>	0.083	0.085	0.080	0.080
39	<i>Coniferous</i>	0.049	0.047	0.061	0.061
40	<i>Deciduous</i>	0.080	0.079	0.057	0.057
41	<i>Mixed</i>	0.134	0.088	0.083	0.083
-	-	RMSE	0.132	0.077	0.077
-	-	NRMSE	0.773	0.453	0.453

Table F.6 – Index of crown differentiation comparison between field and RS derived

Site no.	Class	Field calc. CSDI-4 index	DR and Hyper. CSDI-4 index	FW and Hyper. CSDI-4 index	Best models CSDI-4 index
22	<i>Deciduous</i>	-	-	-	-
23	<i>Coniferous</i>	7.498	3.400	3.310	3.344
24	<i>Deciduous</i>	8.628	4.322	4.507	4.438
25	<i>Deciduous</i>	5.049	5.102	5.142	5.106
26	<i>Deciduous</i>	4.902	3.541	3.988	3.573
27	<i>Deciduous</i>	6.363	6.611	6.821	6.573
28	<i>Coniferous</i>	3.555	3.393	3.328	3.346
29	<i>Deciduous</i>	6.615	3.340	3.316	3.319
30	<i>Coniferous</i>	-	-	-	-
31	<i>Coniferous</i>	2.894	5.169	5.139	5.195
32	<i>Mixed</i>	3.922	3.551	3.523	3.527
33	<i>Coniferous</i>	5.650	3.267	3.167	3.237
34	<i>Mixed</i>	4.900	2.888	2.652	2.756
35	<i>Coniferous</i>	5.102	3.147	3.143	3.118
36	<i>Coniferous</i>	3.194	2.415	2.339	2.361
37	<i>Deciduous</i>	4.327	3.620	3.586	3.575
38	<i>Mixed</i>	4.431	4.580	4.552	4.552
39	<i>Coniferous</i>	3.583	4.367	4.259	4.270
40	<i>Deciduous</i>	3.849	4.114	4.074	4.059
41	<i>Mixed</i>	5.416	4.370	4.260	4.271
-	-	RMSE	1.969	1.971	1.978
-	-	NRMSE	0.343	0.344	0.345

F.2.4 The target and accumulative scoring technique

The scoring method put forward by Cantarello and Newton (2006) requires the assessment of seventeen compositional and structural metrics against an ‘ideal’ target value. Site metrics were defined through statistical means, the results of which were presented in section 8.4. Table F.7 summarizes the remote sensing derived index values for all 20 validation sites against the field based equivalent. A breakdown of the components which make up the score for the field based index is presented in section 4.1.3. The three test datasets produced very similar RMSE and NRMSE values. Of the three test datasets, the combined FW and hyperspectral produced the smallest RMSE and NRMSE values, 2.46 and 0.25 respectively, whereas the highest values were found for the DR and hyperspectral dataset with values of 2.52 and 0.25, again the ‘best’ model subset was in the middle of the other two. Smaller field index values (<7) are overestimated in the case of plot numbers 23, 24 and 38, whereas plot 35 is underestimated. The majority of sites’ field index value matched or were very close to remote sensing derived indices.

Table F.7 – Comparison between the field and remote sensing derived score based index

Site no.	Class	Field calc. score index	DR and Hyper. score index	FW and Hyper. score index	Best models score index
22	<i>Deciduous</i>	9	9	9	9
23	<i>Coniferous</i>	3	6	5	7
24	<i>Deciduous</i>	5	11	8	10
25	<i>Deciduous</i>	10	10	8	9
26	<i>Deciduous</i>	9	10	10	10
27	<i>Deciduous</i>	13	12	12	13
28	<i>Coniferous</i>	6	10	6	8
29	<i>Deciduous</i>	8	10	9	9
30	<i>Coniferous</i>	8	8	6	6
31	<i>Coniferous</i>	6	8	5	7
32	<i>Mixed</i>	11	11	6	7
33	<i>Coniferous</i>	7	5	5	5
34	<i>Mixed</i>	10	11	5	9
35	<i>Coniferous</i>	10	6	6	5
36	<i>Coniferous</i>	8	9	7	9
37	<i>Deciduous</i>	7	10	11	11
38	<i>Mixed</i>	5	9	6	7
39	<i>Coniferous</i>	9	7	7	7
40	<i>Deciduous</i>	8	10	6	7
41	<i>Mixed</i>	10	9	10	11
-	-	RMSE	2.520	2.460	2.510
-	-	NRMSE	0.252	0.246	0.251

In order to test the individual components of the index it is necessary to view a breakdown of each of the three datasets into the individual scoring, illustrated in Tables F.8 to F.10, where differences between field and RS index components were noted. Table F.8 decomposes the scores for the combined DR LiDAR and Hyperspectral models. Plots 24 and 28 performed the worst where 9 of 17 indicators performed sufficiently. Plots 22, 33, and 26 performed the best with 15-16 of 17 correct scores. The average of correct scores was 12. In terms of indicators rather than plots, indicator numbers 3, 6, 7, 9, 15 and 17 were 95-100% correct. Indicator numbers 1, 2, 10, 12, 13 and 16 were only 50% correct.

Table F.9 decomposes the scores for the combined FW LiDAR and Hyperspectral models. Plots 32 and 34 performed the worst in terms for field and RS correspondence, where 10 of 17 indicators were correct. Plots 22, 26, 33 and 38 performed the best with 15-16 out of 17 correct scores. The average correct score was 13. In terms of indicators rather than plots, indicator numbers 3, 7, 9, 14, 15, and 17 were 95-100% correct. Whereas indicator numbers 10, 13 and 16 were only 45-50% correct.

Finally, Table F.10 decomposes the scores for the combined 'Best' models. Plot 24 has the lowest correspondence where only 10 of 17 indicators were correct. Plots 23, 28, 30, 32 and 40 had 11 correspondences. Plots 22, 26, 33 and 38 performed the best with 15-16 out of 17 correct scores. The average correct score was 12.6. In terms of indicators rather than plots, indicator numbers 7, 9, 14, 15 and 17 were 95-100% correct. Indicator numbers 2, 13, and 16 were only 50% correct.

Table F.8 – Decomposition of the combined DR and hyperspectral dataset score index

*Differences between estimated and field recorded are indicated in yellow.

Indicator	site 22	site 23	site 24	site 25	site 26	site 27	site 28	site 29	site 30	site 31	site 32	site 33	site 34	site 35	site 36	site 37	site 38	site 39	site 40	site 41
1. No. of trees (ha-1)	0	1	1	1	0	0	0	0	1	1	0	1	0	1	0	0	1	0	0	0
2. Shannon index for native trees (H)	1	1	1	1	1	1	1	0	1	1	1	0	1	0	1	1	1	0	1	0
3. Basal Area (m ² ha-1)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
4. Mean diameter (cm)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5. Standard dev of diameters (cm)	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1
6. Percentage of big trees	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7. Mean Height (m)	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8. No. of total saplings (ha-1)	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	1	0	1
9. No. of native saplings (ha-1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10. Vol. downed dead wood (m ³ ha-1)	1	1	1	1	1	1	1	1	0	1	1	0	1	0	1	1	0	0	1	1
11. Downed dead wood decay class	1	1	1	1	1	1	1	1	0	1	1	0	1	0	1	1	0	0	1	1
12. Volume of snag (m ³ ha-1)	1	0	0	1	1	1	1	1	1	0	1	0	1	0	1	0	1	1	1	0
13. Snag Decay class	1	0	0	1	1	1	1	1	1	0	1	0	1	0	1	0	1	1	1	1
14. No. of total seedlings (ha-1)	0	0	1	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0
15. No. of native seedlings (ha-1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16. Shannon index for native seedlings (H)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17. No. of ground veg. species	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Score	9	6	11	10	10	12	10	10	8	8	11	5	11	6	9	10	9	7	10	9

Table F.9 – Decomposition of the combined FW and hyperspectral dataset score index

*Differences between estimated and field recorded are indicated in yellow.

Indicator	site 22	site 23	site 24	site 25	site 26	site 27	site 28	site 29	site 30	site 31	site 32	site 33	site 34	site 35	site 36	site 37	site 38	site 39	site 40	site 41
1. No. of trees (ha-1)	0	1	1	1	1	0	1	0	1	1	0	1	1	1	1	1	1	0	0	1
2. Shannon index for native trees (ii)	1	0	0	0	0	1	1	1	1	0	0	0	0	1	0	1	0	0	0	0
3. Basal Area (m2 ha-1)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4. Mean diameter (cm)	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
5. Standard dev. of diameters (cm)	1	0	1	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1
6. Percentage of big trees	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7. Mean Height (m)	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8. No. of total saplings (ha-1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9. No. of native saplings (ha-1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10. Vol. downed dead wood (m3ha-1)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11. Downed dead wood decay class	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12. Volume of snag (m3 ha-1)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13. Snag Decay class	1	0	0	0	0	1	1	1	1	0	0	0	0	0	0	1	0	1	0	1
14. No. of total seedlings (ha-1)	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15. No. of native seedlings (ha-1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16. Shannon index for native seedlings (ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17. No. of ground veg. species	9	5	8	8	8	10	12	6	9	6	5	6	5	5	6	7	11	6	7	10
Score																				

Table F.10 – Decomposition of the combined ‘best’ dataset score index

*Differences between estimated and field recorded are indicated in yellow.

Indicator	site 22	site 23	site 24	site 25	site 26	site 27	site 28	site 29	site 30	site 31	site 32	site 33	site 34	site 35	site 36	site 37	site 38	site 39	site 40	site 41
1. No. of trees (ha ⁻¹)	0	1	1	1	0	1	0	1	1	1	0	1	1	1	1	1	1	0	0	1
2. Shannon index for native trees (ii)	1	1	1	1	1	0	1	0	1	1	1	0	1	0	1	1	1	0	1	0
3. Basal Area (m ² ha ⁻¹)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
4. Mean diameter (cm)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5. Standard dev. of diameters (cm)	1	0	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	1	0	1
6. Percentage of big trees	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7. Mean height (m)	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8. No. of total saplings (ha ⁻¹)	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
9. No. of native saplings (ha ⁻¹)	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
10. Vol. downed dead wood (m ³ ha ⁻¹)	1	1	1	1	1	1	0	1	0	1	1	0	1	0	1	1	1	0	1	1
11. Downed dead wood decay class	1	1	1	1	1	1	0	1	0	1	1	0	1	0	1	1	1	0	1	1
12. Volume of snag (m ³ ha ⁻¹)	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	1
13. Snag Decay class	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	0	1
14. No. of total seedlings (ha ⁻¹)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15. No. of native seedlings (ha ⁻¹)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16. Shannon index for native seedlings (ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17. No. of ground veg. species	9	7	10	9	10	13	8	9	6	7	7	5	9	5	9	11	7	7	7	11